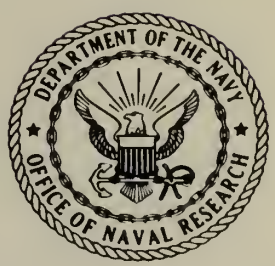
N465

DEPOSITORY
1 APR 1977

NAVAL RESEARCH LOGISTICS QUARTERLY

APR 22 1977
CALIF 93940

MARCH 1977
VOL. 24, NO. 1



OFFICE OF NAVAL RESEARCH

NAVSO P-1278

407-B

NAVAL RESEARCH LOGISTICS QUARTERLY

EDITORS

Murray A. Geisler
Logistics Management Institute

W. H. Marlow
The George Washington University

Bruce J. McDonald
Office of Naval Research

MANAGING EDITOR

Seymour M. Selig
Office of Naval Research
Arlington, Virginia 22217

ASSOCIATE EDITORS

Marvin Denicoff
Office of Naval Research

Alan J. Hoffman
IBM Corporation

Neal D. Glassman
Office of Naval Research

Jack Laderman
Bronx, New York

Thomas L. Saaty
University of Pennsylvania

Henry Solomon
The George Washington University

The Naval Research Logistics Quarterly is devoted to the dissemination of scientific information in logistics and will publish research and expository papers, including those in certain areas of mathematics, statistics, and economics, relevant to the over-all effort to improve the efficiency and effectiveness of logistics operations.

Information for Contributors is indicated on inside back cover.

The Naval Research Logistics Quarterly is published by the Office of Naval Research in the months of March, June, September, and December and can be purchased from the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402. Subscription Price: $11.15 a year in the U.S. and Canada, $13.95 elsewhere. Cost of individual issues may be obtained from the Superintendent of Documents.

The views and opinions expressed in this Journal are those of the authors and not necessarily those of the Office of Naval Research.

Issuance of this periodical approved in accordance with Department of the Navy Publications and Printing Regulations, P-35 (Revised 1-74).

A TWO-ECHELON INVENTORY MODEL WITH PURCHASES, DISPOSITIONS, SHIPMENTS, RETURNS AND TRANSSHIPMENTS

Bruce Hoadley and Daniel P. Heyman

Bell Telephone Laboratories
Holmdel, New Jersey

ABSTRACT

This paper presents a one-period two-echelon inventory model with one warehouse in the first echelon and n warehouses in the second echelon. At the beginning of the period the stock levels at all facilities are adjusted by purchasing or disposing of items at the first echelon, returning or shipping items between the echelons and transshipping items within the second echelon. During the period, demands (which may be negative) are placed on all warehouses in the second echelon and an attempt is made to satisfy shortages either by an expedited shipment from the first echelon to the second echelon or an expedited transshipment within the second echelon. The decision problem is to choose an initial stock level at the first echelon (by a purchase or a disposition) and an initial allocation so as to minimize the initial stock movement costs during the period plus inventory carrying costs and system shortage costs at the end of the period.

It is shown that the objective function takes on one of four forms, depending on the relative magnitudes of the various shipping costs. All four forms of the objective function are derived and proven to be convex.

Several applications of this general model are considered. We also consider multi-period extensions of the general model and an important special case is solved explicitly.

1. INTRODUCTION

In telephone and many military supply systems, when an item is taken out of inventory and placed with a customer, the equipment is still owned by the supplying organizations. For example, customer station apparatus (e.g., telephone sets) are owned by the operating telephone company and aircraft engines or on-board electronic equipment are owned by a branch of the military service. These items are used by the customer and then returned for possible repair and reuse by another customer. Typically, these items are supplied from a multi-echelon multi-location inventory system.

Issues that must be resolved at various levels in a supply line are how much new material should be ordered or how much existing stock should be disposed, and how should the existing stock be allocated to the various echelons. This typically complex problem is further complicated by the fact that at many locations material returns from a lower echelon can exceed demands,

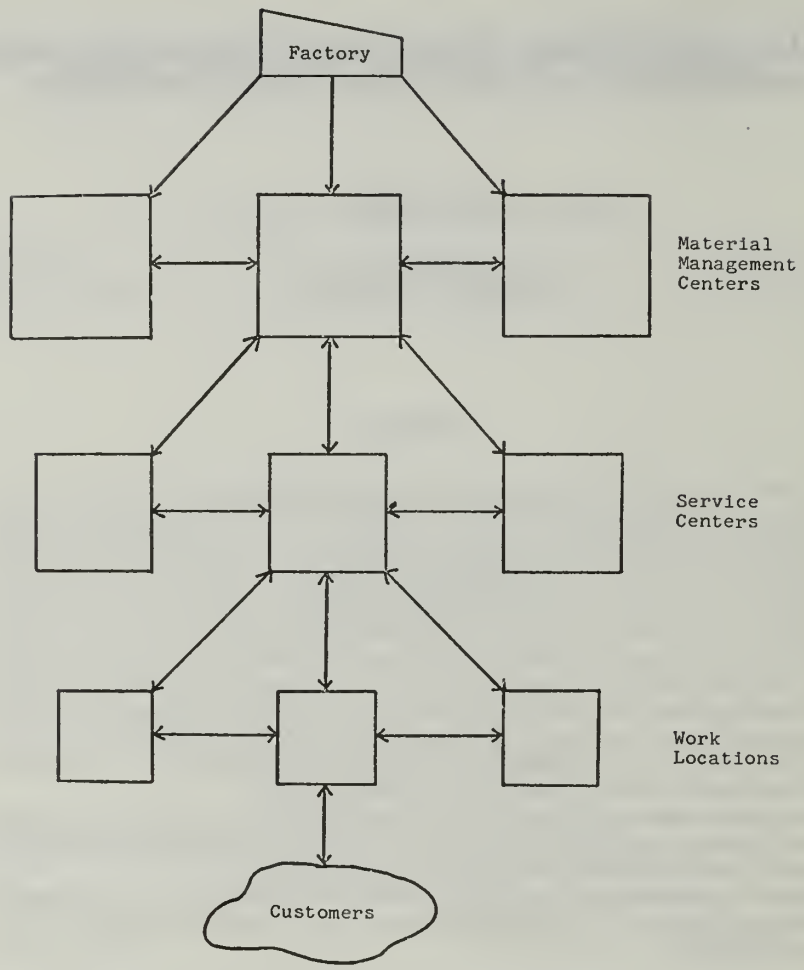


FIGURE 1.—A general supply line.

thus causing a negative net demand. Figure 1 is a schematic diagram of a general supply line where the boxes represent the stocking locations and the arrows represent material flows. The output from the factories must be allocated to a few very large warehouses we will call *material management centers*. Each material management center supplies a set of *service centers* or central stocks which in turn supply the work locations and receive returned material from them. Finally, the material flows from the work locations to the customers and after use the material flows back to the work locations.

The kind of questions that arise in this system are:

(i) How should the factory output be allocated?

(ii) When and how much should one echelon order from another?

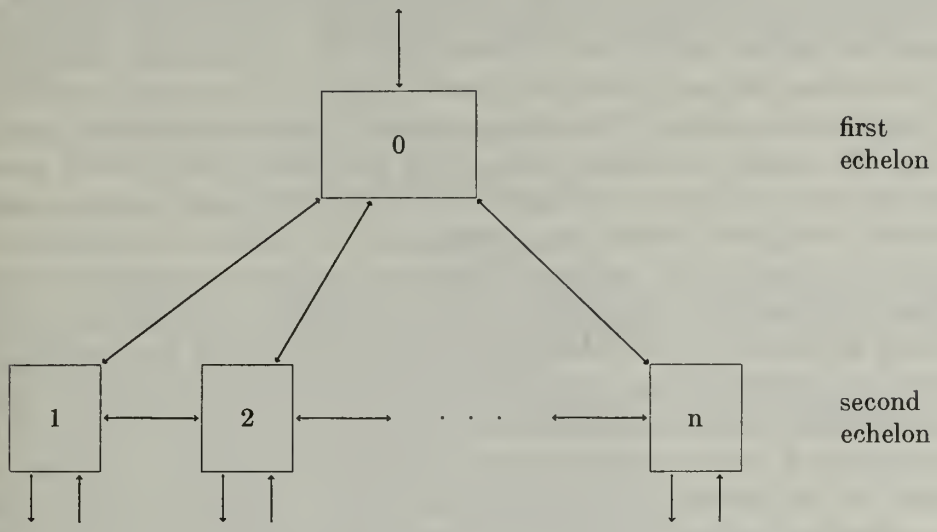
(iii) Should the material be stocked at the material management center which then would supply the service centers on demand, or should some of the material be stocked at the service centers?

(iv) Should the work locations keep stock, or should it be returned to the service centers?

(v) When is it economical to transship material between locations in the same echelon?

1.1. Outline and Motivation

To study these questions, we will consider the two echelon system shown below:



The location in the first echelon will be called facility 0 and the locations in the second echelon will be called facilities 1, 2, . . ., n. Shipments from an outside supplier are received only at facility 0, and external demands and returns occur only at facilities 1, 2, . . ., n. We assume that the net demands (gross demands minus returns) are independent from location to location and from time period to time period. We emphasize that net demand at a second echelon facility may be negative.

Regular and expedited shipments are allowed between any pair of facilities. Since there is no external demand at facility 0, by assuming that it is always less expensive to expedite a shipment between any two second echelon facilities than to expedite shipments between them via facility zero, we do, in effect, prohibit expedited shipments from the second echelon to the first. Facility 0 may dispose of stock but the other facilities can not.

In Section 2 we present a general theory for such a system. In particular, we consider the following decision problems. How much inventory should

a) facility 0 order or dispose;

b) facility 0 ship to facilities 1, 2, . . ., n;

c) facility 0 recall from facilities 1, 2, . . ., n;

d) be transshipped between second echelon facilities. We will also obtain the operating characteristics of the expedited shipments and transshipments.

In Section 3 it is shown how the general model can be applied to three applications that motivated this study. They are the *stock control problem* in a regional material distribution system, the *returns problem* and the *transshipment problem* in a inventory management system.

Multi-period extensions of the general model are discussed in Section 4 and a multi-period formulation of the stock control problem is solved.

1. 2. Relation to Previous Work

Clark [3] is a general survey of multi-echelon systems. Allen [1] considers transshipments among several locations (which correspond to our lower echelon) and finds an algorithm to minimize transshipment plus storage costs. Bessler and Veinott [2] consider a multi-period model of a multi-echelon system with regular and expedited redistributions, but they strongly restrict the flows in their model by assuming that each location can receive shipments from exactly one other location. They obtain bounds on the amount of stock on hand after the regular movements at each location. Clark and Scarf [4] consider purchasing policies for a multi-period multi-echelon system. They solved the problem when activities are arranged in series. They discuss the problem treated in this paper of activities arranged in parallel; they specifically [p. 486] exclude transshipments, claiming they are not done in practice; and state that "The theoretical and computational aspects of the problem become quite complex." Fukuda [9] extends the (activities in series) model of Clark and Scarf by allowing excess stock in each echelon to be disposed of. Gross [8] describes the special case of our single-period model where only regular transshipments are allowed and obtains an algorithm to find the optimal order quantity at each second echelon facility when there is no *a priori* effective bound on the number of items available. Simpson [12] includes such a bound but does not allow transshipments. Das [5] considers a two-location single-echelon single-period model where transshipments are allowed at a given time epoch in the period; returns are not allowed. He finds conditions for the cost function to be convex and for no transshipments to be optimal. The single-echelon model with expedited transshipments of Krishnan and Rao [10] is a very special case of our model.

In our opinion, existing multi-echelon inventory theories often fall short of practical usefulness because they are overly complicated or contain very restrictive assumptions. In this paper we attempt to develop a reasonably simple model that when used with care can help analyze some of the questions raised above.

2. GENERAL FORMULATION AND RESULTS

We assume that events and costs occur in the following manner. At the start of the period, the stocks at each facility are observed. The inventory manager is then given the option of adjusting the stock levels at all facilities in the system. The adjustment can be accomplished by (i) purchasing items at facility 0, (ii) disposing of items at facility 0, (iii) returning items from the second echelon to the first echelon, (iv) shipping items from the first echelon to the second echelon, (v) transshipping between facilities in the second echelon. A per unit cost is associated with each of these transactions, and we assume that the adjustments chosen all occur instantaneously.

After the new stock levels are set, the demands at each second echelon facility are realized. These demands may be negative, indicating that the number of withdrawals requested was smaller than the number of items returned. If the stock at a facility is less than the demand, the manager will attempt to satisfy the demand by making an expedited shipment from the first echelon or an expedited transshipment from another facility. A different holding charge for each echelon is incurred for items remaining in the system at the end of the period. The same holding charge applies to all second echelon facilities. A system shortage cost is incurred for each demand not satisfied by the system. This cost is the same for all facilities and is proportional to the number of unsatisfied demands.

2.1. Notations and Definitions

Let x_i be the stock level at facility $i(i=0, 1, \ldots, n)$ at the beginning of the period (a negative x_i can be interpreted as a number of backorders), $y_i \geq 0$ be the stock level at facility i after the initial adjustments, and D_i be the independent (random) demands at facility $i(i \neq 0)$. Now define the quantities

$$\text{IN}_i=(y_i-x_i)^+ \tag{1}$$

$$\text{OUT}_i=(x_i-y_i)^+ \tag{2}$$

$$\text{V}_i=(D_i-y_i)^+ \tag{3}$$

for $i=1, 2, \ldots, n$, where $(a)^+=\max(a, 0)$. Then at facility i, IN_i is the number of items initially added the stock, OUT_i is the number of items initially subtracted from the stock, and V_i is the shortage before the expedited adjustments are made.

Now we consider the various types of stock movements. At the beginning of the period, let

P=total number of purchases at facility 0,

J=total number of dispositions (junks, sales) at facility 0,

S=total number of shipments from 0 to the second echelon,

T=Total number of transshipments among second echelon facilities,

R=Total number of returns from second echelon to facility 0.

These movements will be called regular movements and the per unit costs of each, for all pairs of facilities, are C_P, C_J, C_S, C_T and C_R, respectively.

For expedited adjustments, let

S^*=total number of expedited shipments from 0 to second echelon,

T^*=total number of expedited transshipments among facilities in the second echelon.

The per unit costs of these movements, for all pairs of facilities, are C^*_S and C^*_T, respectively. Since S^* and T^* depend on the demands, they are random variables.

At the end of the period, after the expedited movements have occurred, let

V_0=system shortage (the excess of total demand over total stock),

s=system shortage cost per unit short,

I_0=inventory at the first echelon,

h_0=carrying cost per unit of I_0,

I=inventory at the second echelon,

h=carrying cost per unit of I.

Since V_0, I_0 and I depend on the demands, they are also random variables.

We note that all movement costs depend only on the number of items moved and not on the distance between the facilities involved. Also there are no fixed costs in the model.

Finally, we define the cost of regular movements by

$$\text{CRM}=C_P \cdot P+C_J \cdot J+C_S \cdot S+C_T \cdot T+C_R \cdot R, \tag{4}$$

the expected cost of expedited movements by

$$\text{CEM}=C^*_S \cdot E(S^*)+C^*_T \cdot E(T^*), \tag{5}$$

the expected carrying cost by

$$\text{CC}=h_0\cdot E(I_0)+h\cdot E(I), \tag{6}$$

the expected system shortage cost by

$$\text{CSS}=s\cdot E(V_0) \tag{7}$$

and the total expected cost by the sum of the four costs given above. Our problem is to choose nonnegative values of $y_0, y_1, y_2, \ldots, y_n$ to minimize the total expected costs.

2.2. General Results for Inventory Levels and Costs

To simplify the formulas, let

$$X=x_0+\sum_{i=1}^{n} x_i$$

and

$$Y=y_0+\sum_{i=1}^{n} y_i.$$

By using a conservation argument, one easily establishes that

$$I_0+I=\left[Y-\sum_{i=1}^{n} D_i\right]^+ \tag{8}$$

$$V_0=\left[\sum_{i=1}^{n} D_i-Y\right]^+. \tag{9}$$

We can also use conservation arguments to show

$$\sum_{i=1}^{n} \text{IN}_i=S+T, \tag{10a}$$

$$\sum_{i=1}^{n} \text{OUT}_i=R+T, \tag{10b}$$

and

$$\sum_{1}^{n} y_i=\sum_{1}^{n} x_i+S-R \tag{10c}$$

or

$$R=\sum_{i=1}^{n}(x_i-y_i)+S. \tag{11}$$

Substituting (10a) and (11) in (4) we obtain

$$\text{CRM}=C_P(Y-X)^+ +C_J(X-Y)^+ +C_R\sum_{i=1}^{n}(x_i-y_i)+C_T\sum_{i=1}^{n}(y_i-x_i)^+ +(C_S+C_R-C_T)\cdot S. \tag{12}$$

Since second echelon shortages are satisfied by either S^*, T^* or V_0, we obtain

$$\sum_{i=1}^{n} V_i=S^*+T^*+V_0. \tag{13}$$

Substituting (13) into (5) yields

$$\text{CEM}=C^*_S E(S^*)+C^*_T E\left[\sum_{i=1}^{n} V_i - S^* - V_0\right],$$

which, using (9), can be written as

$$\text{CEM}=C^*_T E\left[\sum_{i=1}^{n} V_i - V_0\right]+(C^*_S - C^*_T)E[S^*]. \tag{14}$$

Since

$$I_0 = Y - \sum_{i=1}^{n} y_i - S^*, \tag{15}$$

we can obtain from (8)

$$CC=hE[I+I_0]+(h_0-h)E[I_0]=hE\left[Y-\sum_{i=1}^{n} D_i\right]^+ +(h-h_0)\left(E[S^*]+\sum_{i=1}^{n} y_i - Y\right). \tag{16}$$

From (12), (14) and (16) we can derive that the objective function F can be written as

$$\begin{aligned} F=&-hE\left[\sum_{i=1}^{n} D_i\right]+C_R\sum_{i=1}^{n} x_i+C_J X+(h-h_0-C_R)\sum_{i=1}^{n} y_i+(h_0-C_J)Y+C_T\sum_{i=1}^{n}(y_i-x_i)^+ \\ &+(C_P+C_J)(Y-X)^+ +C^*_T E\left[\sum_{i=1}^{n}(D_i-y_i)^+\right]+(h+s-C^*_T)\,E\left[\sum_{i=1}^{n} D_i - Y\right]^+ \\ &+(h-h_0+C^*_S-C^*_T)E(S^*)+(C_S+C_R-C_I)S. \end{aligned} \tag{17}$$

Equation (17) can be written more conveniently by introducing the following functions of $y=(y_0, y_1, y_2, \ldots, y_n)$:

$$f_1(y)=\sum_{i=1}^{n} y_i, \tag{18a}$$

$$f_2(y)=\sum_{i=1}^{n}(y_i-x_i)^+, \tag{18b}$$

$$f_3(y)=\left[\sum_{i=1}^{n}(y_i-x_i)\right]^+, \tag{18c}$$

$$f_4(y)=E\left[\sum_{i=1}^{n}(D_i-y_i)^+\right], \tag{18d}$$

$$f_5(y)=E\left[\sum_{i=1}^{n}(D_i-y_i)\right]^+, \tag{18e}$$

$$f_6(y)=E\left\{\sum_{i=1}^{n}(D_i-y_i)^+ - y_0\right\}^+, \tag{18f}$$

$$f_7(y)=\sum_{i=0}^{n} y_i, \tag{18g}$$

$$f_8(y)=\left(\sum_{i=0}^{n} y_i - X\right)^+, \tag{18h}$$

(18i) $$f_9(y)=E\left[\sum_{i=1}^{n} D_i-\sum_{i=0}^{n} y_i\right]^+.$$

With these definitions, (17) becomes

(19) $$F(y)=-hE\left[\sum_{i=1}^{n} D_i\right]+C_R\sum_{i=1}^{n} x_i+C_JX+(h-h_0-C_R)f_1(y)+(h_0-C_J)f_7(y)+C_Tf_2(y)$$
$$+(C_P+C_J)f_8(y)+C^*{}_Tf_4(y)+(h+s-C^*{}_T)f_9(y)+(h-h_0+C^*{}_S-C^*{}_T)E(S^*)+(C_S+C_R-C_T)S.$$

Recognizing that S and $E(S^*)$ depend on y, we are lead to a study of movement strategies.

2.3. General Results for Movement Strategies

The movement strategies will obviously depend on the movement costs. We distinguish four cases which are analyzed separately.

CASE 1: $C_T>C_R+C_S$

In this case, it is more expensive to transship items between facilities in the second echelon than to first return them to the first echelon and then ship them to another facility in the second echelon. Hence, all the initial stock increases in the second echelon will be effected by a shipment, so

(20) $$S=\sum_{i=1}^{n}(y_i-x_i)^+=f_2(y).$$

CASE 2: $C_T\leq C_R+C_S$

In this case shipments will only be used to satisfy those initial stock increases that cannot be satisfied by transshipments, so

(21) $$S=\left[\sum_{i=1}^{n} y_i-\sum_{i=1}^{n} x_i\right]^+=f_3(y).$$

CASE 3: $C^*{}_T>C^*{}_S$

In this case expedited shipments will be preferred to expedited transshipments, so as many shortages as possible will be satisfied with the stock at the first echelon. Thus,

$$S^*=\text{Min}\left[\sum_{i=1}^{n} V_i, y_0\right]=\sum_{i=1}^{n} V_i-\left[\sum_{i=1}^{n} V_i-y_0\right]^+;$$

and so

(22) $$E(S^*)=f_4(y)-f_6(y).$$

CASE 4: $C^*{}_T\leq C^*{}_S$

In this case expedited transshipments will be used to satisfy as much of the shortages as possible, so

$$T^*=\text{Min}\left[\sum_{i=1}^{n}(D_i-y_i)^+, \sum_{i=1}^{n}(y_i-D_i)^+\right].$$

Since

$$\left(\sum_{i=1}^{n} D_i-\sum_{i=1}^{n} y_i\right)^+$$

is the demand that cannot be satisfied by the initial allocations or by expedited transshipments, we must have

$$\left(\sum_{i=1}^{n} D_i-\sum_{i=1}^{n} y_i\right)^+=S^*+V_0;$$

therefore, from (9)

$$E(S^*)=f_5(y)-f_9(y). \tag{23}$$

Thus, given the relationships among the movement costs, equations (20)–(23) allow us to write (19) in terms of the functions defined in (18), yielding an objective function that is expressed in terms of the decision variables and the distributions of the random demands.

2.4. The Four General Forms of the Objective Function

The four comparisons between movement costs given in the last section give rise to four different cases of the general problem. They are

PROBLEM I: $C^*_T \leq C^*_S$, $C_T > C_R + C_S$,
PROBLEM II: $C^*_T \leq C^*_S$, $C_T \leq C_R + C_S$,
PROBLEM III: $C^*_T > C^*_S$, $C_T > C_R + C_S$,
PROBLEM IV: $C^*_T > C^*_S$, $C_T \leq C_R + C_S$.

Using (20)–(23) the objective function (dropping the constant) for each problem is given in Table 1 below.

TABLE 1.—*Coefficients of the Objective Function*

	Problem			
	I	II	III	IV
$f_k(y)$	$C_T^* \leq C_S^*$ $C_T > C_R + C_S$	$C_T^* \leq C_S^*$ $C_T \leq C_R + C_S$	$C_T^* > C_S^*$ $C_T > C_R + C_S$	$C_T^* > C_S^*$ $C_T \leq C_R + C_S$
$f_1(y)$	$h - h_0 - C_R$	$h - h_0 - C_R$	$h - h_0 - C_R$	$h - h_0 - C_R$
$f_2(y)$	$C_S + C_R$	C_T	$C_S + C_R$	C_T
$f_3(y)$	0	$C_S + C_R - C_T$	0	$C_S + C_R - C_T$
$f_4(y)$	C_T^*	C_T^*	$h - h_0 + C_S^*$	$h - h_0 + C_S^*$
$f_5(y)$	$h - h_0 + C_S^* - C_T^*$	$h - h_0 + C_S^* - C_T^*$	0	0
$f_6(y)$	0	0	$h_0 - h + C_T^* - C_S^*$	$h_0 - h + C_T^* - C_S^*$
$f_7(y)$	$h_0 - C_J$	$h_0 - C_J$	$h_0 - C_J$	$h_0 - C_J$
$f_8(y)$	$C_P + C_J$	$C_P + C_J$	$C_P + C_J$	$C_P + C_J$
$f_9(y)$	$h_0 + s - C_S^*$	$h_0 + s - C_S^*$	$h + s - C_T^*$	$h + s - C_T^*$

2.5. Convexity of the Objective Function

We will now give some simple sufficient conditions for the objective functions shown in Table 1 to convex. First, we show that $f_i(y)$, $i=1, 2, \ldots, 9$ are convex functions.

THEOREM 1: The functions $f_i(y)$, $i=1, 2, \ldots, 9$ are convex functions of y.

PROOF: a) Since $f_1(y)$ is a linear function, both $-f_1(y)$ and $f_1(y)$ are convex. b) The function $(y_i-x_i)^+$ is obviously a convex function of y_i, so $f_2(y)$ is convex because it is a sum of convex functions. c) $f_3(y)$ is the maximum of

$$\sum_{i=1}^{n}(y_i-x_i)$$

and the null function, both of which are convex, so $f_3(y)$ is convex. d) Since

$$\sum_{i=1}^{n}(D_i-y_i)^+$$

is clearly convex in y for any choice of D_i, $i=1, 2, \ldots, n$, $f_4(y)$ is convex because convexity is preserved under expectations. e) Since

$$\left[\sum_{i=1}^{n}(D_i-y_i)\right]^+$$

is convex for any choice of D_i, $i=1, \ldots, n$, $f_5(y)$ is convex. Since

$$\left\{\sum_{i=1}^{n}(D_i-y_i)^+-y_0\right\}^+$$

is the maximum of a convex function and the null function for any choice of D_i, $i=1, \ldots, n$, $f_6(y)$ is convex. The remainder of the proof follows in a similar manner. Q.E.D.

So sufficient conditions for the objective functions to be convex are that the coefficients of all the f_k's except f_1 and f_7 are nonnegative. A special case of interest for which the above is true is

$$h=h_0 \tag{24a}$$

$$C_P \geq -C_J \tag{24b}$$

$$s \geq C^*_S \tag{24c}$$

$$s \geq C^*_T; \tag{24d}$$

i.e., the carrying cost factor is the same for both echelons, the purchase price is greater than the disposition value and the system shortage cost is greater than both the expedited shipping and transshipment costs.

If the inequalities in (24) hold, then the problem of finding a nonnegative value of y to minimize (19) is a convex programming problem. This type of mathematical programming problem has been extensively studied (see, e.g., Eaves and Saigal [6], Fiacco and McCormick [7], Mifflin [11], and Zangwill [13]) and algorithms and computer codes are available to solve it. We should point out that the functions $f_k(y)$ are not necessarily differentiable. For example, $f_2(y)$ is not differentiable at those points where $y_i=x_i$, $i=1, \ldots, n$ and $f_4(y)$ is not differentiable at those points where $y_i=d_i$ if $\Pr\{D_i=d_i\}>0$.

However, the manner in which the derivative fails to exist is very simple, so algorithms that require a differentiable objective function should be adequate for this problem. For the inventory applications considered here, y must be composed of integers to be physically meaningful. Since we do not know of an algorithm that can solve such a problem, we have chosen to ignore the integer aspects of y. Thus, the usual caveats about solving integer problems as noninteger ones apply.

3. APPLICATIONS

In this section we discuss how the general model can be used in several applications and find the optimal solution for some of them. To keep the characterizations relatively simple, we assume that the c.d.f. of D_i is of the form

$$\Pr\{D_i \leq y\} = \begin{cases} \int_{-\infty}^{y} g_i(x)dx & y<0 \\ \int_{-\infty}^{0} g_i(x)dx + p_i & y=0 \\ \int_{-\infty}^{0} g_i(x)dx + p_i + \int_{0}^{y} f_i(x)dx & y>0; \end{cases} \tag{25}$$

i.e.,

$$\Pr\{D_i=0\} = p_i$$

but otherwise D_i is a continuous random variable with a density function. This assumption allows us to model both high and low volume products.

In deriving the characterizations, the following Lemmas are useful.

LEMMA 1: For a random variable D,

$$\frac{\partial^+}{\partial y} E[D-y]^+ = -Pr\{D>y\}$$

$$\frac{\partial^-}{\partial y} E[D-y]^+ = -Pr\{D\geq y\},$$

where ∂^+ and ∂^- denote right and left derivatives respectively.

PROOF: The Lemma follows directly by interchanging the derivative and expectation operators. Q.E.D.

The second Lemma is due to Allen [1, Theorem 1, p. 339] and we will use it extensively.

LEMMA 2: (Allen) Let $F(y)$ be a real, convex differentiable function in n-dimensional Euclidean space. Let $l_i(u_i)$ be a minimum (maximum) of the i-th coordinate y_i if y_i is bounded from below (above). Let R be the set of all points so restricted. Then a necessary and sufficient condition that a point y^0 maximize F over R is that either

$$\frac{\partial F(y^0)}{\partial y_i} = 0$$

or

$$y_i^0 = l_i \text{ and } \frac{\partial F(y^0)}{\partial y_i} > 0$$

or

$$y_i = u_i \text{ and } \frac{\partial F(y^0)}{\partial y_i} < 0$$

is satisfied for $i=1, 2, \ldots, n$.

Allen also describes an algorithm for finding y^0, so a method for obtaining numerical solutions already exists.

3.1. The Stock Control Problem

Consider a regional distribution system consisting of a material management center (MMC) (the first echelon) and several service centers (the second echelon). The stocks owned by the region are purchased through the MMC from the factory and are allocated to all facilities in the region. In the case of a product where growth is taking place at all service centers, it is appropriate to assume that $\Pr\{D_i \geq 0\}=1$ and that there is no need to return material to the MMC, transship material between service centers or to dispose of material on a regular basis; hence, we also assume that $y_i \geq x_i$, $i=1, \ldots, n$. Also, since regular shipments are cheaper than expedited shipments, and the average distance between service centers is less than the average distance between the MMC and the service centers, we assume that $C_S \leq C^*_T \leq C^*_S \leq s$. This is a special case of either Problem I or II which are the same in this case because $f_2(y)=f_3(y)$. To avoid unnecessary complication, it is also assumed that $h=h_0$.

Note that the assumption $C^*_T \leq C^*_S$ amounts to assuming that when a demand cannot be satisfied by the local service center, an expedited transshipment is made from another service center if possible.

The multiperiod extension of this problem is solved in Section 4.

From the general theory, the problem is to find a vector

$$y^0=(y_0^0, y_1^0, \ldots, y_n^0)$$

which minimizes

$$(26)\quad F_1(y)=C_P\left(\sum_{i=0}^{n} y_i - X\right)+h\left[\sum_{i=0}^{n} y-\sum_{i=1}^{n} E(D_i)\right]+(h+s-C^*_S)E\left[\sum_{i=1}^{n} D_i-\sum_{i=0}^{n} y_i\right]^+$$

$$+C_S\sum_{i=1}^{n}(y_i-x_i)+C^*_T E\left[\sum_{i=1}^{n}(D_i-y_i)^+\right]+(C^*_S-C^*_T)E\left[\sum_{i=1}^{n}(D_i-y_i)\right]^+,$$

subject to

$$y_0 \geq 0$$

$$y_i \geq x_i \qquad i=1, \ldots, n$$

$$(27)\qquad \sum_{i=0}^{n} y_i \geq X.$$

According to Lemma 2, to find y^0, we need only construct a vector for which either

$$(28a)\qquad \frac{\partial F_1(y^0)}{\partial y_0}=0 \text{ or } y_0^0=0 \text{ and } \frac{\partial^+ F_1(y^0)}{\partial y_0}>0$$

and for which either

$$(28b)\qquad \frac{\partial F_1(y^0)}{\partial y_i}=0 \text{ or } y_i^0=x_i \text{ and } \frac{\partial^+ F_1(y^0)}{\partial y_i}>0$$

for $i=1, \ldots, n$, and for which

$$(28c)\qquad \sum_{i=1}^{n} y^0{}_i \geq X.$$

To construct such a vector, first note that

$$\frac{\partial F_1(y)}{\partial y_0}=(>)\,0 \text{ iff}$$

(29a) $$\Pr\left\{\sum_{i=1}^{n} D_i > \sum_{i=0}^{n} y_i\right\} = (<) \frac{C_P+h}{h+s-C^*_S},$$

and

$$\frac{\partial F_1(y)}{\partial y_i} = (>)0 \qquad i=1, \ldots, n \text{ iff}$$

(29b) $$\Pr\{D_i > y_i\} = (<)G(y),$$

where

$$G(y)=\left[C_P+h-(h+s-C^*_S)\Pr\left\{\sum_{i=1}^{n} D_i > \sum_{i=0}^{n} y_i\right\}+C_S-(C^*_S-C^*_T)\Pr\left\{\sum_{i=1}^{n} D_i > \sum_{i=1}^{n} y_i\right\}\right]\Big/ C^*_T.$$

To keep the analysis relatively simple, we assume that $s > C_P + C_S^*$ so that

$$(C_P+h)/(h+s-C^*_S)<1,$$

and assume that $\overline{Y}$ defined by

(30a) $$\Pr\left\{\sum_{i=1}^{n} D_i > \overline{Y}\right\} = \frac{C_P+h}{h+s-C^*_S}$$

is greater than X.

Now for $0 \leq \epsilon \leq 1$ and $i=1, \ldots, n$, define

(30b) $$y_i(\epsilon) = \begin{cases} x_i & \text{if } \Pr\{D_i > x_i\} < \epsilon \\ \text{a solution to } \Pr\{D_i > y_i\} = \epsilon & \text{otherwise,} \end{cases}$$

and in the latter case $y_i(\epsilon) > x_i$, and define

(30c) $$y_0(\epsilon) = \left[\overline{Y} - \sum_{i=1}^{n} y_i(\epsilon)\right]^+,$$

$$R(\epsilon) = G(y(\epsilon)).$$

Note that either

$$\overline{Y} = \sum_{i=0}^{n} y_i(\epsilon)$$

or $y_0(\epsilon)=0$ and

$$\overline{Y} < \sum_{i=0}^{n} y_i(\epsilon);$$

hence, $y(\epsilon)$ satisfies (28a) and (28c) for any ϵ.

Now,

$$R(0) = \frac{C_P+h+C_S}{C^*_T},$$

$$R(1) = \frac{C_S-(C^*_S-C^*_T)\Pr\left\{\sum_{i=1}^{n} D_i > \sum_{i=1}^{n} x_i\right\}}{C^*_T} < 1,$$

and $R(\epsilon)$ is a nonincreasing function of ϵ; therefore, $R(\epsilon)=\epsilon$ has a solution, ϵ^*, and $y^0=y(\epsilon^*)$ satisfies (28).

To interpret the results, about the only thing that can be said in general is that all second echelon facilities that receive stock will experience the same shortage probability, and if some stock is initially allocated to the first echelon, then the system shortage probability is

$$[C_P+h]/[h+sC^*_s].$$

The first conclusion is the same as a result of Simpson [12, Theorem 2, p. 801] and extends that result to the more general situation where transshipments and expedited movements are possible.

3.2. The Returns Problem

A problem that arises when net demand at some second echelon facilities may be negative is: how many items should these facilities return to the first echelon? We shall call this the *returns problem.*

The returns problem is the special case of Problem III where $C_T>C_R+C_S$ (hence $T\equiv 0$) and $C_S=C_S^*<C_T^*$ (hence regular shipments will be zero). To avoid unnecessary complications, we assume that $h=h_0$ and that

$$\sum_{i=0}^{n} y_i$$

is given and equal to

$$\sum_{i=0}^{n} x_i.$$

Thus we seek a vector $y=(y_1, \ldots, y_n)$ to minimize

(31)

$$F_2(y)=-C_R\sum_{i=1}^{n} y_i+C^*_S E\left[\sum_{i=1}^{n}(D_i-y_i)^+\right]+(C^*_T-C^*_S)E\left[\sum_{i=1}^{n}(D_i-y_i)^+-x_0-\sum_{i=1}^{n}(x_i-y_i)\right]^+,$$

subject to

(32) $$0\leq y_i\leq x_i,\ i=1, 2, \ldots, n.$$

An algorithm for solving exactly this type of problem is given in [11]. We shall restrict ourselves to characterizing the optimal solution.

Since F_2 is a convex function, Lemma 2 asserts that y^0 is an optimal point if and only if either

(33a) $$\frac{\partial F_2(y^0)}{\partial y_i}=0,\ 0<y_i^0<x_i,$$

or

(33b) $$\frac{\partial^+ F_2(y^0)}{\partial y_i}>0 \text{ and } y_i^0=0,$$

or

(33c) $$\frac{\partial^- F_2(y^0)}{\partial y_i}<0 \text{ and } y_i^0=x_i,$$

holds for $i=1, 2, \ldots, n$.

By interchanging the derivative and expectation operators we get

$$\text{(34a)}\quad \frac{\partial^+}{\partial y_i} E\left[\sum_{i=1}^{n}(D_i-y_i)^+-x_0-\sum_{i=1}^{n}(x_i-y_i)\right]^+ = E\,\frac{\partial^+}{\partial y_i}\left[\sum_{i=1}^{n}(D_i-y_i)^+-x_0-\sum_{i=1}^{n}(x_i-y_i)\right]^+$$

$$= \Pr\left\{D_i \le y_i,\ \sum_{i=1}^{n}(D_i-y_i)^+ \ge y_0\right\},$$

$$\text{(34b)}\quad \frac{\partial^-}{\partial y_i} E\left[\sum_{i=1}^{n}(D_i-y_i)^+-x_0-\sum_{i=1}^{n}(x_i-y_i)\right]^+ = \Pr\left\{D_i < y_i,\ \sum_{i=1}^{n}(D_i-y_i)^+ > y_0\right\},$$

where

$$\frac{\partial^+}{\partial y_i} = \frac{\partial^-}{\partial y_i}$$

whenever $0<y_i<x_i$.

Recalling that $V_i=(D_i-y_i)^+$ and

$$R=\sum_{i=1}^{n}(x_i-y_i)$$

and combining (33) and (34) yields: y^0 is an optimal solution if either

(35a)

$$C^*_T \Pr\left\{V_i=0,\ \sum_{i=1}^{n} V_i \ge x_0+R\right\} = C_R+C^*_S\left[\Pr\{V_i>0\}+\Pr\left\{V_i=0,\ \sum_{i=1}^{n} V_i \ge x_0+R\right\}\right], \quad 0<y_i{}^0<x_i$$

or

(35b)

$$C^*_T \Pr\left\{D_i\le 0,\ \sum_{i=1}^{n} V_i \ge x_0+R\right\} > C_R+C^*_S\left[\Pr\{D_i>0\}+\Pr\left\{D_i\le 0,\ \sum_{i=1}^{n} V_i \ge x_0+R\right\}\right], \qquad y_i{}^0=0$$

or

(35c)

$$C^*_T \Pr\left\{D_i<x_i,\ \sum_{i=1}^{n} V_i > x_0+R\right\} < C_R+C^*_S\left[\Pr\{D_i\ge x_i\}+\Pr\left\{D_i<x_i,\ \sum_{i=1}^{n} V_i > x_0+R\right\}\right], \qquad y_i{}^0=x_i.$$

To help interpret (35) note that the right side of (35a) is the expected additional return plus expedited shipment cost associated with an additional unit returned from facility i to facility 0, and the left side is the expected expedited transshipment cost of that same unit if it were not returned.

3.3. The Transshipment Problem

Consider a multi-service center company where orders placed on a service center can be satisfied by an expedited transshipment from another service center when the original service center is out of stock. For this system we are interested in determining if transshipments done in anticipation of demand are economical. For this problem there is no first echelon and the only types of regular and expedited movements are transshipments. From the general theory, it can be shown that the optimization problem is to find a vector $\underline{y}=(y_1, \ldots, y_n)$ which minimizes

$$F_3(\underline{y})=C_T\sum_{i=1}^{n}(y_i-x_i)^+ + C^*_T E\left[\sum_{i=1}^{n}(D_i-y_i)^+\right] \tag{36}$$

subject to

$$\sum_{i=1}^{n} y_i=\sum_{i=1}^{n} x_i. \tag{37}$$

The solution, $\underline{y}^0$, to this problem can be obtained by minimizing the Lagrangian function

$$L(\underline{y}, \lambda)=F_3(\underline{y})-\lambda\left[\sum_{i=1}^{n} x_i-\sum_{i=1}^{n} y_i\right], \tag{38}$$

subject to $\underline{y}\geq 0$, $\lambda\geq 0$,

$$\sum_{i=1}^{n} y_i=\sum_{i=1}^{n} x_i.$$

According to Lemma 2 we need only construct $\underline{y}^0$, λ^0 for which either

$$\frac{\partial L(\underline{y}^0, \lambda^0)}{\partial \underline{y}_i}=0 \qquad 0<y_i^0<x_i \qquad x_i<y_i^0<\sum_{j=1}^{n} x_j \tag{39a}$$

or

$$\frac{\partial^+ L(\underline{y}^0, \lambda^0)}{\partial y_i}\geq 0 \qquad y_i^0=0 \tag{39b}$$

or

$$\frac{\partial^- L(\underline{y}^0, \lambda^0)}{\partial y_i}\leq 0\leq\frac{\partial^+ L(\underline{y}^0, \lambda^0)}{\partial y_i} \qquad y_i^0=x_i \tag{39c}$$

or

$$\frac{\partial^- L(\underline{y}^0, \lambda^0)}{\partial y_i}\leq 0 \qquad y_i^0=\sum_{j=1}^{n} x_j. \tag{39d}$$

These conditions imply

$$Pr\{D_i\geq y_i^0\}=\frac{\lambda^0}{C^*_T} \qquad 0<y_i^0<x_i \tag{40a}$$

or

$$Pr\{D_i\geq y_i^0\}=\frac{\lambda^0+C_T}{C^*_T} \qquad x_i<y_i^0<\sum_{j=1}^{n} x_i \tag{40b}$$

or

$$Pr\{D_i\geq 0\}\leq\frac{\lambda^0}{C^*_T} \qquad y_i^0=0 \tag{40c}$$

or

$$\frac{\lambda^0}{C^*_T}\leq Pr\{D_i\geq x_i\}\leq\frac{\lambda^0+C_T}{C^*_T} \qquad y_i^0=x_i \tag{40d}$$

or

$$\frac{C_T+\lambda^0}{C^*_T}\leq Pr\{D_i\geq y_i^0\} \qquad y_i^0=\sum_{j=1}^{n}x_j. \tag{40e}$$

The Lagrange multiplier λ^0 can be interpreted as the "cost" of using up one of the items available for transshipment. Rewrite (40b) as

$$C^*_T \Pr\{D_i\geq y_i^0\}=\lambda^0+C_T$$

and observe that for the last item received at i, the left side is the expected savings in expedited transshipments and the right side is the cost of placing them at locations i. For those locations which tranship some of their stock, the same interpretation holds with $C_T=0$ (because the cost of transshipping to themselves is zero) which yields (40a). So all facilities which receive transshipments have the same shortage probability, which is larger than the common shortage probability for all locations which transship some but not all of their stock.

Another conclusion that follows from these conditions is that if

$$\Pr\{D_i\geq x_i\}\leq\frac{C_T}{C^*_T}$$

for all i, then $\lambda^0=0$, $y^0=(x_1, \ldots, x_n)$; i.e., nothing is transshipped.

4. MULTIPERIOD EXTENSION

The multiperiod extensions of our model using either the discounted total expected cost criterion or the expected cost per period criterion are straightforward to formulate and in general difficult to solve. However, there is one situation where a formulation of the multiperiod problem with the expected cost per period criterion is tractable. This is when the solution to the one period problem does not depend on $x=(x_0, x_1, \ldots, x_n)$; because then it is reasonable in the multiperiod problem (it may even be optimal) to restrict attention to policies that do not depend on x.

The policy is specified by a vector $y\geq 0$ and the cost rate for an infinite horizon can be in terms of y, the vector of inventory levels just before the demands occur, and then minimized to find the optimal value. To illustrate this we consider the infinite horizon extension of the stock control problem.

Using the policy y let $C(y)$ be the asymptotic cost rate, $L(y)$ be the expected holding plus shortage cost rate for the system, $P(y)$ be the expected cost rate of purchases, $M(y)$ be the expected cost rate for regular movements and $M^*(y)$ be the expected cost rate for expedited movements. We assume that all demand processes are stationary and all cost factors are constant with time. The assumptions made in section 3.1 are invoked here also.

Clearly

$$C(y)=L(y)+P(y)+M(y)+M^*(y). \tag{41}$$

From the nature of the policy we find

$$L(y)=sE\left(\sum_{i=1}^{n}D_i-\sum_{i=0}^{n}y_i\right)^+ +hE\left(\sum_{i=0}^{n}y_i-\sum_{i=1}^{n}D_i\right)^+, \tag{42a}$$

and

$$P(y)=C_P\sum_{i=1}^{n}E(D_i) \tag{42b}$$

immediately. Since the only regular movements are shipments and in each period the number shipped to the second echelon will be just enough to replace the stock depletion from demand in the last period,

$$M(y)=C_SE\left[\text{Min}\left(\sum_{i=1}^{n}D_i,\ \sum_{i=1}^{n}y_i\right)\right]=C_SE\left\{\sum_{i=1}^{n}D_i-\left[\sum_{i=1}^{n}(D_i-y_i)\right]^+\right\}. \tag{42c}$$

Since $C^*_T\leq C_S^*$ we obtain from (9), (14) and (23)

$$\begin{aligned}M^*(y)&=C^*_TE\left(\sum_{i=1}^{n}V_i-V_0\right)+(C^*_S-C^*_T)E(S^*)\\&=C^*_TE\left[\sum_{i=1}^{n}(D_i-y_i)^+-\left(\sum_{i=1}^{n}D_i-\sum_{i=0}^{n}y_i\right)^+\right]\\&\qquad+(C^*_S-C^*_T)\left\{E\left[\sum_{i=1}^{n}(D_i-y_i)\right]^+-E\left[\sum_{i=1}^{n}D_i-\sum_{i=0}^{n}y_i\right]^+\right\}.\end{aligned} \tag{42d}$$

Substituting (42a)–(42d) into (41) and collecting terms yields

$$\begin{aligned}C(y)&=(C_P+C_S-h)\sum_{i=1}^{n}E(D_i)+h\sum_{i=0}^{n}y_i+\left(h+S-C^*_S\right)E\left[\sum_{i=1}^{n}D_i-\sum_{i=0}^{n}y_i\right]^+\\&\qquad+C^*_TE\left[\sum_{i=1}^{n}(D_i-y_i)^+\right]\end{aligned}$$

$$+(C^*_S-C^*_T-C_S)E\left[\sum_{i=1}^{n}(D_i-y_i)\right]^+. \tag{43}$$

Again according to Lemma 2, to find y^0, we need only construct a vector for which either

$$\frac{\partial C(y^0)}{\partial y_i}=0 \text{ or } y_i{}^0=0 \text{ and } \frac{\partial^+C(y^0)}{\partial y_i}>0, \tag{44}$$

for $i=0, 1, \ldots, n$. The vector y^0 may be constructed by using the methods in section 3.1; we omit the details.

The solution is different from the single period solution, but has the same characteristics. One of these characteristics is that the probability of shortage at a second echelon facility can be quite large. This is because once an item has been shipped to the second echelon, the only way to move it to a different second echelon facility is by an expensive expedited transshipment. When the item is retained at the first echelon and is not needed during the period in question, it may be subsequently shipped by the cheaper mode to a more needy second echelon facility.

Unfortunately, the above approach does not work in general. For example, the one period solutions of the returns problem and the transshipment problem discussed in Section 3.2 and 3.3 both depend on x and therefore the analyses cannot readily be extended to the multiperiod case.

5. ACKNOWLEDGMENT

The authors would like to acknowledge the contributions made to this work by Alan Rolfe, who formulated a version of the returns problem which helped lead to the general problem presented here.

REFERENCES

[1] Allen, S. G., "Redistribution of Total Stock over Several User Locations," Naval Research Logistics Quarterly, *5*, 51–59 (1958).

[2] Bessler, S., and A. F. Veinott Jr., "Optimal Policy for a Dynamic Multi-Echelon Inventory Problem," Naval Research Logistics Quarterly, *13*, 355–389 (1966).

[3] Clark, A. J., "An Informal Survey of Multi-Echelon Inventory Theory," Naval Research Logistics Quarterly, *19*, 621–650 (1972).

[4] ——— and H. Scarf, "Optimal Policies for a Multi-Echelon Inventory Problem," Management Science, *6*, 475–490 (1960).

[5] Das, Chandrasekhar, "Supply and Redistribution Rules for Two-Location Inventory Systems: One-Period Analysis" Management Science, *21*, 765–776 (1975).

[6] Eaves, B. C., and R. Saigal, "Homotopies for Computing Fixed Points in Unbounded Regions," Mathematical Programming, *3*, 225–237 (1972).

[7] Fiacco, A. V., and G. P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques* (Wiley, New York, 1968).

[8] Gross, D., "Centralized Inventory Control in Multilocation Supply Systems," Chapter 3 in *Multistage Inventory Models and Techniques*, H. E. Scarf et al. eds. (Stanford University Press, Stanford, California, 1963).

[9] Fukuda, Y., "Optimal Disposal Policies," Naval Research Logistics Quarterly, *8*, 221–227 (1961).

[10] Krishnan, K. S., and V. R. K. Rao, "Inventory Control in N Warehouses," The Journal of Industrial Engineering, *16*, 212–215 (1965).

[11] Mifflin, Robert, "A Nonderivative Algorithm for Minimization of a Function of Several Bounded Variables," paper presented at the joint O.R.S.A. T.I.M.S. Meeting (October, 1974).

[12] Simpson, K. E., Jr., "A Theory of Allocation of Stocks to Warehouses," Operations Research, *7*, 797–805 (1959).

[13] Zangwill, W. I., *Nonlinear Programming* (Prentice-Hall, Englewood Cliffs, 1969).

OPTIMAL REJECT ALLOWANCE WITH CONSTANT MARGINAL PRODUCTION EFFICIENCY‡

Avraham Beja

The Leon Recanati Graduate School of Business Administration
Tel-Aviv University
Tel-Aviv, Israel

ABSTRACT

A job shop must fulfill an order for N good items. Production is conducted in "lots," and the number of good items in a lot can be accurately determined only after production of that lot is completed. If the number of good items falls short of the outstanding order, the shop must produce further lots, as necessary.

Processes with "constant marginal production efficiency" are investigated. The revealed structure allows efficient exact computation of optimal policy. The resulting minimal cost exhibits a consistent (but not universal) pattern whereby higher quality of production is advantageous even at proportionately higher marginal cost.

I. INTRODUCTION AND SUMMARY

Consider a job shop with an outstanding order for N items. Production is conducted in "runs" or "lots"* involving a set up cost and direct production cost. Some of the items in any lot may be defective, and the number of good items can be accurately determined only after the production run is completed. If the number of good items produced falls short of the outstanding order, the shop must initiate further runs as necessary. The operational problem, then, is to determine the optimal lot size that minimizes total costs incurred to fulfill the order.

We tend to expect that under "reasonable" conditions the optimal lot size will be larger than the outstanding order to allow for the anticipated rejects, hence the traditional terminology of "reject allowance" or "shrinkage allowance" associated with this problem.†

Our analysis of the reject allowance problem includes questions of structure and of computation. The study of structure concerns the qualitative properties (monotonicity, unimodality) of the relevant functions such as the minimal cost and the optimal policy. Computationally, the focus

‡This study was supported in part by the Israel Institute for Business Research, University Campus, Ramat Aviv, Tel Aviv, Israel.

*The terms run and lot will be used interchangeably.

†Some conditions for this to be indeed true will become clear through subsequent results.

is on efficient methods for calculating the optimal decision in any given problem. The two apsects are naturally interrelated; understanding the structure allows more efficient computation, and computational algorithms provide insight on structure.

The computational aspect of the reject allowance problem has attracted considerable interest (e.g. [2, 3, 4, 5, 7, 8, 9]). The analysis usually tended to concentrate on the search for approximations to the cost function that would facilitate the computation. The suggested ("optimal") production decision would then be the point that satisfies the usual first order conditions for the surrogate (or approximate) cost function. Before a solution based on an analysis of this kind can be safely employed, however, some quantitative measures of the degree of approximation must be established. Otherwise, there is a very little assurance that the suggested solution is in any meaningful sense "close" to the optimal lot size or the minimal cost. Unfortunately, no such measures are given in the literature. For example, Hillier [5] suggests an approximation for the first differences of the cost function in a special production process. Even if the approximation is assymptotically valid, no general assessment of the error is available for finite order sizes. Some writers (e.g. [2, 7, 8]) assume that all costs beyond that of the first production run and one eventual additional set up cost may be ignored. This is naturally restricted to cases with relatively high set-up cost, but even then it is not at all clear that production runs beyond the first one should be either few in number or small in size. Generally, the basic difficulty with all these approximations is that the analysis of the structure of optimal policy is based not only on assumptions regarding the problem's parameters, but also on *assumptions regarding the optimal policy.*

The present study is devoted to the analysis of a class of production processes characterized by a property called "constant marginal efficiency." This includes as a special case the widely considered process with constant marginal production cost and binomially distributed defectives (e.g. [2, 3, 5, 7, 8, 9]).

The computational aspects of the problem are addressed first. Here, we seek efficient methods for an *exact* solution to the problem. The model is formally presented in Section II, and it is shown that, in principle, an exact solution can be computed recursively by enumeration of all "relevant" lot sizes. It is then suggested that if a certain structure of the model is established, the range of computation will be drastically reduced. To achieve this, a Markovian decision model formulation of the problem is presented and analyzed in Section III. The direct computational implications are drawn in Section IV. Roughly speaking, we establish that (i) there is an optimal policy whereby lot sizes are strictly increasing with the outstanding order, and (ii) if for some outstanding order a lot of $n+1$ is inferior to a lot of n, larger lots need not be considered. Given these results, the necessary volume of computation is not only "reasonable," but indeed trivial with modern computing capacity even for fairly large order sizes. In Section V, we investigate the sensitivity of total cost to the quality of production. The results show that under a fairly wide range of circumstances (though not always) higher quality is advantageous even at a proportionately higher marginal production cost. Some remarks on possible extensions and applications conclude the study in Section VI.

II. A STRAIGHTFORWARD FORMULATION OF THE MODEL

Production is defined by a stochastic process $\{X_j, j=1, 2 \ldots\}$ where $X_j=1$ if the j^{th} item in the lot is good and $X_j=0$ if it is defective.

$$G_k=\sum_{j=1}^{k} X_j$$

denotes the number of good items in a lot of size k. For $j=1, 2 \ldots$ let $q_j=P[X_j=1]$ be a given technological property of the process, assumed independent of $X_1, X_2, \ldots X_{j-1}$. This property may be equivalently represented by

$$\left.\begin{aligned} &q(j, k)=P[G_k=j] \\ &Q(j, k)=P[G_k \geq j]=\sum_{i=j}^{k} q(i, k) \end{aligned}\right| \begin{aligned} &k=1, 2, \ldots \\ &j=0, 1, \ldots, k \end{aligned}$$

or by

where by elementary probability theory we have for $k>1$

$$\left.\begin{aligned} q(j, k)&=q_k q(j-1, k-1)+(1-q_k)q(j, k-1) \qquad \text{for } j=0, \ldots k \\ &=0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \text{otherwise} \end{aligned}\right\} \tag{2.1}$$

The production cost for a lot of size k is

$$Y(k)=C_o+C_1+\ldots+C_k$$

where $C_o \geq 0$ is a set-up cost and $C_j>0$ is the expected direct (marginal) cost for producing the j^{th} item in a run. $h_j=C_j/q_j$ is therefore the expected cost per good item produced at the j^{th} "*stage*" in a production run, and $e_j=1/h_j$ can therefore be considered the production "*efficiency*" at the j^{th} stage. It is naturally assumed throughout that $Y(k)$ is unbounded.

We define three inter-related cost functions:

$F(N)$ is the (minimal) total expected cost to fulfill an outstanding order of N items when an optimal policy is carried throughout.

$F(N, n)$ is the expected cost to fulfill an outstanding order of N items, if the first lot is of size n and *all* subsuequent lots are optimal.

$f(N, n)$ is the expected cost to fulfill an outstanding order of N items if a lot of size n is produced whenever the outstanding order is of size N and an optimal lot size is produced whenever the outstanding order is less than N.

The three functions satisfy the following relationships:

$$f(N, n)=Y(n)+q(0, n)f(N, n)+\sum_{j=1}^{N-1} q(j, n)F(N-j) \tag{2.2}$$

$$f(N, n)=\left\{Y(n)+\sum_{j=1}^{N-1} q(j, n)F(N-j)\right\} \Big/ \{1-q(0, n)\} \tag{2.3}$$

$$F(N, n)=Y(n)+\sum_{j=0}^{N-1} q(j, n)F(N-j) \tag{2.4}$$

$$F(N)=\min_{n=1, 2, \ldots} F(N, n)=\min_{n=1, 2, \ldots} f(N, n) \tag{2.5}$$

$F(N)$ clearly exists, and can in principle be computed in a finite number of steps as follows:
Start with $N=1$. Compute $f(1, n)=Y(n)/\{1-q(0, n)\}$ as in (2.3) for $n=1, 2, \ldots$ until, say, m where

$$Y(m) \geq \min_{n=1, \ldots, m-1} f(1, n)=\hat{F}(N)$$

Clearly $f(1, n) > \hat{F}(N)$ for $n > m$, and hence $F(N) = \hat{F}(N)$. Proceed similarly with $N = 2, 3, \ldots$, at each stage using in (2.3) the values $F(M)$ $M = 1, \ldots, N-1$ computed previously.

The practical difficulty with this procedure originates from the wide range of values of n for which $f(N, n)$ must be evaluated at each N. The present study helps to reduce this range for processes with constant marginal production efficiency, i.e. $h_j = h$ for all j. This is achieved by establishing two important properties of these processes:

(i) If a lot of n is optimal for an outstanding order of N, then for an order of $N+1$ there is an optimal lot $r > n$ (monotonicity of the optimal policy).

(ii) If n is optimal for N and $f(N, k+1) \geq f(N, k)$ then $n \leq k$.

These properties are proved indirectly by an alternative Markovian decision model formulation of the same process.*

III. AN ALTERNATIVE MARKOVIAN FORMULATION

Define a Markov decision model as follows: for $k = 0, 1, 2 \ldots N = 0, 1, 2 \ldots$, the process is said to be in state (k, N) if the outstanding order is N, set up has been achieved and k items have been manufactured but not inspected for defectives. The process operates in discrete time, and at each period the following actions are available: (i) Continue production at an immediate cost of C_{k+1}, with a transition to state $(k+1, N)$ with probability 1. (ii) Stop production and inspect the k items already produced for the number j of good items; if $j < N$ a new production run must be initiated, hence with probability $q(j, k) j = 0, \ldots N-1$, there is a transition to $(0, N-j)$ at a cost C_0, and with probability $Q(N, k)$ the process terminates.†

An optimal policy is a decision rule that minimizes the expected sum of all costs until the process terminates. Let $v(k, N)$ be the minimal expected total cost from the time the process is in state (k, N) until termination. $v(k, N)$ clearly exists, and satisfies

$$v(k, N) = \min \{v^1(k, N), C_{k+1} + v(k+1, N)\} \tag{3.1}$$

where

$$v^1(k, N) = \sum_{j=0}^{N} q(j, k)\{C_0 + v(0, N-j)\} \tag{3.2}$$

$v^1(k, N)$ is the conditional minimal expected cost from state (k, N) until Termination, given that the process is to be inspected in the present period.

LEMMA 3.1: $v^1(k+1, N+1) = (1-q_{k+1})v^1(k, N+1) + q_{k+1}v^1(k, N)$

PROOF: By conditioning the expected cost on X_{k+1}.

Let

$$H(k, N) = v^1(k, N) - v^1(k+1, N).$$ Then for $N = 1, 2, \ldots$

$$H(k, N) = v^1(k, N) - (1-q_{k+1})v^1(k, N) - q_{k+1}v^1(k, N-1)$$

---

*Both properties are very intuitive. Should the reader feel that this calls for a direct proof, he may find it instructive to try proving any of them directly for the simplest case where $C_j = C$ and $q_j = q$ for all j.

†To ensure termination under optimal policy even if $C_0 = 0$, inspection is not allowed if $N > 0$ and $k = 0$.

$$H(k, N)=q_{k+1}\{v^1(k, N)-v^1(k, N-1)\} \tag{3.3}$$

$$H(k+1, N)=(1-q_{k+1})v^1(k, N)+q_{k+1}v^1(k, N-1)-(1-q_{k+2})v^1(k+1, N)-q_{k+2}v^1(k+1, N-1)$$

$$H(k+1, N)=(1-q_{k+2})H(k, N)+q_{k+2}H(k, N-1)-(q_{k+1}-q_{k+2})\{v^1(k, N)-v^1(k, N-1)\}$$

and substituting (3.3)

$$H(k+1, N)=\left\{1-q_{k+2}-\frac{q_{k+1}-q_{k+2}}{q_{k+1}}\right\}H(k, N)+q_{k+2}H(k, N-1)$$

$$H(k+1, N)=\frac{q_{k+2}}{q_{k+1}}\{(1-q_{k+1})H(k, N)+q_{k+1}H(k, N-1)\} \tag{3.4}$$

Let M denote the set of all states in which it is optimal to manufacture another item and I the set of all states in which it is optimal to stop and inspect, i.e.

$$(k, N)\epsilon M \text{ if } v(k, N)=C_{k+1}+v(k+1, N) \tag{3.5}$$

$$(k, N)\epsilon I \text{ if } v(k, N)=v^1(k, N) \tag{3.6}$$

then for $(k, N)\epsilon I$

$$v^1(k, N)\leq C_{k+1}+v(k+1, N)\leq C_{k+1}+v^1(k+1, N)$$

$$H(k, N)\leq C_{k+1} \text{ for all } (k, N)\epsilon I \tag{3.7}$$

and if

$$v^1(k, N)\geq C_{k+1}+v^1(k+1, N)$$

then

$$v^1(k, N)\geq C_{k+1}+v(k+1, N) \text{ and } (k, N)\epsilon M,$$

thus

$$H(k, N)\geq C_{k+1} \text{ implies } (k, N)\epsilon M \tag{3.8}$$

We shall also use the following important property:

LEMMA 3.2: $v(0, N+1)-v(0, N)\geq h$

PROOF: Let $r=$ min $\{k: (k, N+1)\epsilon I\}$ (clearly r exists and $r>0$) Then

$$v(0, N+1)=C_1+ \ . \ . \ . \ +C_r+v^1(r, N+1)$$

$$v(0, N+1)=C_1+ \ . \ . \ . \ +C_r+(1-q_r)v^1(r-1, N+1)+q_rv^1(r-1, N) \tag{3.9}$$

$$v(0, N+1)\geq C_1+ \ . \ . \ . \ +C_{r-1}+C_r+(1-q_r)v(r-1, N+1)+q_rv(r-1, N)$$

but by recursive application of (3.1)

$$v(0, N)\leq C_1+ \ . \ . \ . \ +Cr_{-1}+v(r-1, N)$$

and by assumption $v(0, N+1)=C_1+ \ . \ . \ . \ +C_{r-1}+v(r-1, N+1)$. Hence from (3.9)

$$v(0, N+1)\geq C_r+(1-q_r)v(0, N+1)+q_rv(0, N)$$

$$v(0, N+1)-v(0, N)\geq C_r/q_r=h$$

COROLLARY 3.3: For $N=1, 2, \ldots$ $H(0, N)\geq C_1$

PROOF: By Lemma 3.2 and (3.3).

We now present the basis structure of the reject allowance model.

THEOREM 3.4: There is a strictly increasing sequence $<n^*(N)>$ of non-negative integers, such that in state (k, N) it is optimal to continue production if $k<n^*(N)$ and to stop and inspect if $k\geq n^*(N)$.

PROOF: We show by induction on N that

(i) for $k<n^*(N)$ $(k, N)\epsilon M$ and $H(k, N)\geq C_{k+1}$

(ii) for $k\geq n^*(N)$ $(k, N)\epsilon I$ and $H(k, N)<C_{k+1}$

(iii) $n^*(N+1)>n^*(N)$.

For $N=0$ (i) and (ii) are trivially true with $n^*(0)=0$, because $v^1(k, 0)=0$ and $H(k, 0)=0$ for all k. For $N=1$, let

$$m=\min\{k: H(k, 1)<C_{k+1}\}$$

m exists because if $H(k, 1)\geq C_{k+1}$ for all k then $(k, 1)\epsilon M$ for all k and $v(0, 1)=C_1+\ldots C_n+v(n, 1)$ for all n, which is impossible because $v(0, 1)$ clearly exists and $Y(n)$ is unbounded. By Corollary 3.3 $m>0$. For $k>m$ $H(k, 1)<C_{k+1}$ by induction, because if $H(j, 1)<C_{j+1}$ then by (3.4)

$$H(j+1, 1)<\frac{q_{j+2}}{q_{j+1}}(1-q_{j+1})C_{j+1}\leq\frac{q_{j+2}}{q_{j+1}}C_{j+1}=C_{j+2}.$$

Let $n^*(1)=m$.

Assume inductively that (i) and (ii) are true for N. By Corollary 3.3 $H(0, N+1)\geq C_1$, and we note that if $H(k-1, N)\geq C_k$ and $H(k-1, N+1)\geq C_k$ then

$$H(k, N+1)\geq\frac{q_{k+1}}{q_k}\{(1-q_k)C_k+q_kC_k\}=\frac{q_{k+1}}{q_k}C_k=C_{k+1}$$

hence, by induction on k $H(k, N+1)\geq C_{k+1}$ for $k=1, \ldots, n^*(N)$. Let

$$r=\min\{k: H(k, N+1)<C_{k+1}\}$$

Existence of $v(0, N+1)$ implies existence of r, and clearly $r>n^*(N)$. For $k>r$ $H(k-1, N)<C_k$ and hence, again by induction on k, $H(k, N+1)<C_{k+1}$ because if $H(k-1, N+1)<C_k$ then

$$H(k, N+1)<\frac{q_{k+1}}{q_k}\{(1-q_k)C_k+q_kC_k\}=\frac{q_{k+1}}{q_k}C_k=C_{k+1}.$$

Let $n^*(N+1)=r$, and the theorem is proved by induction for all N.

IV. STRUCTURE AND COMPUTATION

The formulation of Section III is convenient for understanding the structure of the reject allowance model, but not necessarily for computing the optimal policy. It therefore remains to be seen how the results of Theorem 3.4 bear on the direct computation of optimal policy by $f(N, n)$. For $N=1, 2, \ldots$ $F(N)=C_o+v(0, N)$, and by (2.4) and (3.2)

$$v^1(k, N)=F(N, k)-Y(k) \tag{4.1}$$

and if $n^*(N)=m$

$$F(N)=C_o+v(0, N)=C_o+C_1+ \ . \ . \ . \ C_m+v^1(m, N)=F(N, m)$$

The (strict) monotonicity of $n^*(N)$ thus allows the evaluations of $f(N+1, n)$ to start with

$$n=n^*(N)+1.$$

Now consider

$$H(k, N)=\{F(N, k)-Y(k)\}-\{F(N, k+1)-Y(k+1)\}=F(N, k)-F(N, k+1)+C_{k+1}$$

so that $H(k, N)\geq C_{k+1}$ is equivalent to $F(N, k)\geq F(N, k+1)$ and $H(k, N)<C_{k+1}$ is equivalent to $F(N, k)<F(N, k+1)$.

Theorem 3.4 therefore establishes that $F(N, n)$ is "quasi-convex," since $F(N, n+1)\leq F(N, n)$ for $n<n^*(N)$ and $F(N, n+1)>F(N, n)$ for $n\geq n^*(N)$. Computationally, however, the interest is in $f(N, n)$ rather than $F(N, n)$, and one further step is necessary.

THEOREM 4.1: $f(N, n+1)>f(N, n)$ implies $F(N, n+1)>F(N, n)$.

PROOF: By (2.2) and (2.4)

$$f(N, n+1)-f(N, n)=F(N, n+1)+q(0, n+1)f(N, n+1)$$
$$-q(0, n+1)F(N)-F(N, n)-q(0, n)f(N, n)+q(0, n)F(N)$$

$$f(N, n+1)-f(N, n)=F(N, n+1)-F(N, n)+q(0, n+1)\{f(N, n+1)-f(N, n)\}$$
$$-\{q(0, n)-q(0, n+1)\}f(N, n)+\{q(0, n)-q(0, n+1)\}F(N)$$

$$\{q(0, n)-q(0, n+1)\}\{f(N, n)-F(N)\} \tag{4.3}$$
$$+\{1-q(0, n+1)\}\{f(N, n+1)-f(N, n)\}=F(N, n+1)-F(N, n)$$

The first term on the left hand side of (4.3) is clearly non-negative, and if the second term is positive the right hand side must be positive. Q.E.D.

Note also that at $n=n^*(N)$ the right hand side of (4.3) is positive, the first term on the left hand side vanishes, and hence $f(N, n+1)>f(N, n)$.

Theorems 3.4 and 4.1 thus establish a very effective upper bound for the computations of $f(N, n)$, which always stop at $n^*(N)+1$, where $f(N, n)-f(N, n-1)$ first becomes positive. If the values of q_j are not extremely small there may be no need for more than just a few evaluations of $f(N, n)$ for each N. Experimentation suggests that for N as high as 100 the computations need usually take no more than a few seconds with a large scale computer.

V. THE EFFECTS OF QUALITY ON COST

Returning to the general structure of the reject allowance problem, the next question to be considered concerns the relationship between $F(N)$ and the system's parameters. Processes with constant marginal efficiency can be identified by three elements: the set up cost C_o, the marginal efficiency e or, equivalently, its reciprocal value h, and the sequence $\langle q\rangle=q_1, q_2, \ . \ . \ .$ which represents the "*quality*" of production, with a lower percentage of defectives for higher $\langle q\rangle$. Given $\langle q\rangle$

and C_o, $F(N)$ clearly increases with h, and, similarly, given $\langle q \rangle$ and h, $F(N)$ increases with C_o. The question, then, is: given C_o and h, how does $F(N)$ depend on $\langle q \rangle$? If $C_o=0$, the answer is immediate.

THEOREM 5.1: If $C_o=0$ then $F(N)=Nh$ (regardless of $\langle q \rangle$).

PROOF: $F(N) \leqq F(M)+F(N-M)$ for $M<N$, hence $F(N) \leqq NF(1)$. But $F(1) \leqq f(1, 1)$, and when $C_o=0$ then by (2.3) $f(1, 1)=C_1/q_1=h$, hence $F(N) \leqq Nh$. By recursive applictaion of Lemma 3.2 $F(N) \geqq Nh$, and hence $F(N)=Nh$. Q.E.D.

It is easy to see (e.g. by a numerical example such as the one given at the end of this section) that, generally, $F(N)$ need not be independent of $\langle q \rangle$. In this section we investigate the effects of quality on total cost by making various comparisons at different levels of generalization between processes with the same (arbitrary) C_o and h. The comparisons exhibit a consistent pattern whereby lower quality is more costly. As we shall see, however, this statement must be accepted only in the exact sense of the theorems proved below; a careless "universal" interpretation is false.

One way to compare the costs associated with two production processes is by formulating a decision problem where, before each lot is produced, a choice is available as to which of the two processes to use for that lot. If an optimal policy always uses the same process, then certainly this process is (costwise) at least as good as the other. This approach is not only methodologically convenient, but also firmly rooted in the motivation underlying the analysis.

We shall make repeated use of Howard's *policy iteration principle.** For our purposes, this well known principle may be conveniently summarized in rough words as follows: policy I is preferred (strictly preferred) to policy II if, and only if, using policy I *once* and policy II afterwards is preferred (*strictly preferred*) to policy II.

THEOREM 5.2: Let production process I be specified by C_o, h and $\langle q \rangle$, and process II by C_o, h and $\langle p \rangle$, where $p_i=q_i$ for $i \neq n$ and $p_n > \sup\{q_i,\ i=n, n+1, \ldots\}$. Then $F_{II}(N) \leqq F_I(N)$.

PROOF: It is convenient to introduce a hypothetical third process—process 0—specified by C_o, h and $\langle q^0 \rangle$ where $q_n{}^0=0$ and $q_i{}^0=q_i$ for $i \neq n$.† Consider a decision problem involving a choice among the three processes. $F_{II}(0)=F_I(0)=F_o(0)$.

Assume inductively that for $M=0, 1, \ldots N-1$ process II is optimal and

$$F_{II}(M) \leqq F_I(M),\ F_{II}(M) \leqq F_o(M).$$

For N, let a non-stationary policy A use process I for the first lot and process II for all subsequent lots, at a total cost of $F^A(N, k)$ if the size of the first lot is k. Similarly, policy B uses process 0 once and process II afterwards. Let r be optimal for A and m for B, i.e.

$$F^A(N, r)=\min_{k=1, 2, \ldots} F^A(N, k) \tag{5.1}$$

$$F^B(N, m)=\min_{k=1, 2, \ldots} F^B(N, k) \tag{5.2}$$

We shall show that $F_{II}(N) \leqq F^A(N, r)$ and $F_{II}(N) \leqq F^B(N, m)$, so that by Howard's principle process II is optimal for N.

$$F^B(N, m)=Y_o(m)+\sum_{i=0}^{N-1} q_o(i, m)F_{II}(N-i)$$

*Cf. Howard [6] or Blackwell [1].

†This is clearly just a convenient way of assigning the indices in a process which is essentially defined by C_o, h, $\langle w \rangle$, where $w_i=q_i{}^0$ for $i \leqq n-1$ $w_i=q_i+1^0$ for $i \leqq n$, with $w_i>0$ for all i.

so that if $m<n$

$$F^B(N, m)=F_{II}(N, m)\geq F_{II}(N),$$

and if $m=n$

$$F^B(N, m)=F_{II}(N, n-1)\geq F_{II}(N).$$

If $m\geq n+1$, then by (2.1)

$$F^B(N, m)=Y_o(m-1)+q_m h+\sum_{i=0}^{N-1}\{(1-q_m)q_o(i, m-1)+q_m q_o(i-1, m-1)\}F_{II}(N-i)$$

(5.3)

$$F^B(N, m)=F^B(N, m-1)+q_m D(m-1)$$

where

(5.4)
$$D(m-1)=h-\sum_{i=0}^{N-1}\{q_o(i, m-1)-q_o(i-1, m-1)\}F_{II}(N-i)$$

Equation (5.2) implies that $D(m-1)\leq 0$. Consider $F_{II}(N, m-1)$ as defined in (2.4)

$$F_{II}(N, m-1)=Y_{II}(m-1)+\sum_{i=0}^{N-1} q_{II}(i, m-1)F_{II}(N-i)$$

An immediate extension of (2.1), again by elementary probability theory, established that for $k\geq n$

$$q_{II}(i, k)=(1-p_n)q_0(i, k)+p_n q_0(i-1, k)$$

and thus

$$F_{II}(N, m-1)=Y_o(m-1)+p_n h+\sum_{i=0}^{N-1}\{(1-p_n)q_o(i, m-1)+p_n q_o(i-1, m-1)\}F_{II}(N-i)$$

$$F_{II}(N, m-1)=F^B(N, m-1)+p_n D(m-1)$$

and by (5.3)

$$F_{II}(N, m-1)=F^B(N, m)+(p_n-q_m)D(m-1)$$

$D(m-1)\leq 0$ and $p_n\geq q_m$ for $m\geq n+1$ imply that

(5.5)
$$F^B(N, m)\geq F_{II}(N, m-1)\geq F_{II}(N).$$

Now consider

$$F^A(N, r)=Y_I(r)+\sum_{i=0}^{N-1} q_I(i, r)F_{II}(N-i)$$

and if $r\leq n-1$

$$F^A(N, r)=F_{II}(N, r)\geq F_{II}(N).$$

For $r\geq n$

$$F^A(N, r)=Y_o(r)+q_n h+\sum_{i=0}^{N-1}\{(1-q_n)q_o(i, r)+q_n q_o(i-1, r)\}F_{II}(N-i)$$

(5.6)
$$F^A(N, r)=F^B(N, r)+q_n D(r)$$

where $D(r)$ is defined as in (5.4). Also

$$F_{II}(N, r)=Y_o(r)+p_nh+\sum_{i=0}^{N-1}\{(1-p_n)q_o(i, r)+p_nq_o(i-1, r)\}F_I(N-i)$$

$$F_{II}(N, r)=F^B(N, r)+p_nD(r)$$

and by (5.6)

$$F_{II}(N, r)=F^A(N, r)+(p_n-q_n)D(r) \tag{5.7}$$

If $D(r)\leq 0$ then (5.7) and $p_n\geq q_n$ imply $F^A(N, r)\geq F_{II}(N, r)\geq F_{II}(N)$, and if $D(r)\geq 0$ then (5.6), (5.2) and (5.5) imply $F^A(N, r)\geq F^B(N, r)\geq F^B(N, m)\geq F_{II}(N)$. We have thus proved that

$$F_{II}(N)\leq F^A(N, r)$$

as well as $F_{II}(N)\leq F^B(N, m)$, process II is optimal for N, and in particular

$$F_{II}(N)\leq F_I(N),$$

so that the proof is valid by induction for all N.

It has already been established earlier by Lemma 3.2 that an option to obtain a good item at cost h *separated from the production process* (i.e. between two production runs) always pays, since $F(N)\geq F(N-1)+h$. Theorem 5.2 allows an extension of this property, which may perhaps be best formulated immediately in the operational context.

COROLLARY 5.3: An option to obtain a good item at cost h *at any stage of the production process* always pays.

PROOF: Let the basic quality sequence of the process be $\langle w\rangle$. If the option involves perfect production of, say, the n^{th} item, let $\langle q\rangle=\langle w\rangle$ and let $\langle p\rangle$ be defined by $p_n=1$, $p_i=w_i$ for $i\neq n$, and by Theorem 5.2 the option pays.

If the option allows the good item to be obtained between the $n-1$ and the n^{th} item in a run, let $\langle p\rangle$ and $\langle q\rangle$ be defined by $p_i=q_i=w_i$ for $i\leq n-1$, $p_n=1$, $q_n=0$ and $p_i=q_i=w_{i-1}$ for $i\geq n+1$. Theorem 5.2 again holds. Q.E.D.

The next theorem generalizes the "advantage of quality" presented in Theorem 5.2.

THEOREM 5.4. Let process I be specified by C_o, h, and $\langle q\rangle$, and process II by C_o, h, and $\langle p\rangle$. If $q_i\geq q_{i+1}$ and $p_i\geq q_i$ for $i=1, 2, \ldots$ then $F_{II}(N)\leq F_I(N)$.

PROOF: Define a sequence of processes, indexed by $n=2, 3, \ldots$, with C_o, h and $\langle w^n\rangle$ where $w^n{}_i=p_i$ for $i=1 \ldots n-1$ and $w^n{}_i=q_i$ for $i=n, n+1, \ldots$. Then by Theorem 5.2 $F_2(N)\leq F_I(N)$ and $F_{n+1}(N)\leq F_n(N)$ for $n=2, 3, \ldots$. Since the range of lot sizes that are relevant for N is bounded, $\langle p\rangle$ is operationally equivalent to $\langle w^n\rangle$ for n sufficiently large and $F_{II}(N)\leq F_I(N)$ Q.E.D.

Processes with non-increasing $\langle q\rangle$ as in Theorem 5.4 are analogous to the notion of "increasing failure rate" in reliability theory in the sense that negative effects of time and wear dominate production. When effects of "running in" and "learning" are dominant, non-decreasing quality sequences $\langle q\rangle$ are of interest. Here the "advantage of quality" take a slightly different form.

THEOREM 5.5: Let process I be specified by C_o, h, and $\langle q\rangle$, and process II by C_o, h, and $\langle p\rangle$. If $q_i\leq q_{i+1}$ and $p_i=q_{i+n}$ for all i and some (natural) n then $F_{II}(N)\leq F_I(N)$.

PROOF: If suffices to consider $n=1$, because for $n=2, 3, \ldots$ the theorem then follows immediately by induction.

Consider the decision problem involving a choice between the two processes. $F_{II}(0)=F_I(0)$, and we assume inductively that for $M=0, \ldots N-1$ process II is optimal and $F_{II}(M)\leq F_I(M)$. Suppose that process I is optimal for N, and that the corresponding optimal lot size is r. Let policy A use process II for the first lot and an optimal policy for all subsequent lots, with a total cost of $F^A(N, k)$ if the size of the first lot is k. By assumption

$$F(N)=Y_I(r)+\sum_{i=0}^{N-1} q_I(i, r)F(N-i)$$

$$F(N)=Y_{II}(r-1)+q_1h+\sum_{i=0}^{N-1}\{(1-q_1)q_{II}(i, r-1)+q_1q_{II}(i-1, r-1\}F(N-i)$$

$$F(N)=F^A(N, r-1)+q_1D_{II}(r-1) \tag{5.8}$$

where

$$D_{II}(r-1)=h-\sum_{i=0}^{N-1}\{q_{II}(i, r-1)-q_{II}(i-1, r-1)\}F(N-i)$$

Also

$$F^A(N, r)=Y_{II}(r)+\sum_{i=0}^{N-1} q_{II}(i, r)F(N-i)$$

$$F^A(N, r)=Y_{II}(r-1)+q_rh+\sum_{i=0}^{N-1}\{(1-q_r)q_{II}(i, r-1)+q_rq_{II}(i-1, r-1)\}F(N-i)$$

$$F^A(N, r)=F^A(N, r-1)+q_rD_{II}(r-1) \tag{5.9}$$

and by (5.8)

$$F^A(N, r)=F(N)+(q_r-q_1)D_{II}(r-1) \tag{5.10}$$

If $D_{II}(r-1)\geq 0$ then by (5.8) $F^A(N, r-1)\leq F(N)$ and if $D_{II}(r-1)\leq 0$ then (since $q_r\geq q_1$) by (5.10) $F^A(N, r)\leq F(N)$. Hence by Howard's principle process II is optimal for N, $F_{II}(N)\leq F_I(N)$ and the theorem proved by induction on N and on n.

It is tempting to try to extend the fairly general advantage of quality of Theorem 5.4 to arbitrary quality sequences $\langle q\rangle$, or, in view of Theorem 5.5, at least to non-decreasing sequences. This extension is false, however, as demonstrated by the following numerical example.

AN EXAMPLE WITH DISADVANTAGE TO QUALITY:

Let $C_o=h=1$, $\langle q\rangle=0.1, 1, 1, \ldots$, $\langle p\rangle=0.2, 1, 1 \ldots$. The first process involves a lower quality at the first stage of production, but still $F_I(1)=2.1$ whereas for the second "higher quality process" $F_{II}(1)=2.2>F_I(1)$.

VI. EXTENSIONS, APPLICATIONS AND LIMITATIONS

The formulation of (2.2) through (2.5) does not include an explicit "salvage value" for defective items or for good ones produced in excess of the outstanding order. Nevertheless, if items of both kinds have the same salvage value, the process still falls within the scope of our model by a straightforward transformation. Let the index s refer to a process with salvage value S, then

$$F_s(N, n)=Y_s(n)+\sum_{i=0}^{N-1} q(i, n)\{F_s(N-i)-(n-i)S\}-Q(N, n)(n-N)S \tag{6.1}$$

define

$$F_T(N, n)=F_s(N, n)-NS$$

$$Y_T(n)=Y_s(n)-nS \text{ (or equivalently } C_j^T=C_j^s-S)$$

and then (6.1) becomes

$$F_T(N, n)+NS=Y_T(n)+nS+\sum_{i=0}^{N-1} q(i, n)\{F_T(N-i)+(N-i)S-(n-i)S\}-Q(N, n)(n-N)S$$

$$F_T(N, n)+NS=Y_T(n)+nS+\sum_{i=0}^{N-1} q(i, n)F_T(N-i)+\{1-Q(N, n)\}(N-n)S-Q(N, n)(n-N)S$$

$$F_T(N, n)=Y_T(n)+\sum_{i=0}^{N-1} q(i, n)F_T(N-i) \tag{6.2}$$

and (6.2), which is equivalent to (2.4), solves the problem because minimization of $F_T(N, n)$ is clearly equivalent to minimization of $F_s(N, n)$. $Y_T(n)$ is the "excess production cost" of a lot of n items (beyond its salvage value) and it should be emphasized that the transformation is useful only if "constant marginal efficiency" applies to the excess marginal cost C_j^T. $F_T(N)$ is the excess cost of an order for N good items, beyond their salvage value.

In Wadsworth and Chang's model [9] the first set up cost need not equal the cost of subsequent set ups. This variation does not affect our analysis in any way, because the difference in cost between the first and subsequent set ups may simply be charged to $F(N)$, with a constant set up cost considered throughout.

If the technological or economic parameters change between runs, let consecutive runs be indexed by $t=1, 2, \ldots$ so that (3.2) becomes

$$v_t^1(k, N)=\sum_{j=0}^{N-1} q_t(j, k)\{C_0^{t+1}+v_{t+1}(0, N-j)\}$$

Monotonicity of $n_t^*(N)$ and quasi-convexity of $F_t(N, n)$ still hold, because Theorem 3.4 is true for all t, as can be verified by careful inspection of the proofs involved. This interesting structural property has little, if any, computational value, however, because $F_{t+1}(M)$ must be evaluated for $M=0, \ldots, N-1$ *before* $f_t(N, n)$ can be computed.

The unavoidable limitation of the model is, of course, the inherent restriction to processes with constant marginal efficiency. This domain naturally includes the dominantly popular $q_j=q$ and $C_j=c$ for all j. If q_j varies during production, the model still includes all cases where direct cost is proportional to the number of good times produced, rather than the total number of items. It can serve as a fairly good approximation when quality incentive payments are a dominant part of direct production cost, or when a machine's power consumption—although varying with time during production—strongly affects both quality and cost, etc. Within the context of constant marginal efficiency, effects of "running-in" and "learning" (increasing q_j) or "fatigue" and "wear" (decreasing q_j) are certainly allowed in unlimited variations.

An interesting special application concerns contracting. Suppose N good items are needed, and an agreement with a subcontractor can be secured whereby an order for the delivery of n items is

placed under the understanding that only good items are paid for, at constant price h (independent of n). Then if $q(j, n)$ is the probability that out of n items delivered j items are good and C_0 is the cost of placing the order, the model applies.

The study of more general processes, where changes in marginal cost reflect more than changes in q_j, is certainly of interest. It is doubtful, however, that the structure of such processes is nearly as powerful for the computation of an exact solution as in processes with constant marginal efficiency. In particular, it can be readily verified that $F(N, n)$ need not in general be quasi-convex, and a local minimum does not guarantee a global minimum. For processes with decreasing marginal efficiency we do not even expect that $n^*(N) > n^*(N-1)$, nor indeed that $n^*(N) \geq N$.

With constant marginal efficiency $n^*(N) > n^*(N-1)$ insures that $n^*(N) \geq N$, so that at least for that class the term "reject allowance" is justified.

REFERENCES

[1] Blackwell, D., 'Discrete Dynamic Programming," Annals of Mathematical Statistics, *33*, 719–726 (1962).

[2] Bowman, E. H., and R. N. Fetter, *Analysis for Production Management*, Revised Edition (Richard D. Irwin, Inc., Homewood, Illinois, 1960), 324–330.

[3] Goode, H. P., and S. Saltzman, "Computing Optimum Shrinkage Allowances for Small Order Sizes," The Journal of Industrial Engineering, 57–61 (January–February, 1961).

[4] Gregory, W. R., and A. Beged-Dov, "On the Determination of Optimal Shrinkage Allowance in a Job Shop," The Journal of Industrial Engineering (April, 1967).

[5] Hillier, F. S., "Reject Allowances for Job Lot Orders," The Journal of Industrial Engineering, 311–316 (November–December, 1963).

[6] Howard, R. A., *Dynamic Programming and Markov Processes* (John Wiley and Sons, New York, 1960).

[7] Levitan, R. E., 'The Optimum Reject Allowance Problem," Management Science, *6*, 172–186 (1960).

[8] Llewellyn, R. W., "Order Sizes for Job Lot Manufacturing," The Journal of Industrial Engineering, 176–180 (May–June, 1959).

[9] Wadsworth, H. M., and S. H. Chang, "The Reject Allowance Problem: An Analysis and Application to Job Lot Production," The Journal of Industrial Engineering, 127–132 (May–June, 1964).

A CHANCE-CONSTRAINED DISTRIBUTION PROBLEM

Richard M. Reese and Andrew C. Stedry

School of Economics and Management
Oakland University
Rochester, Michigan

ABSTRACT

The transportation model with supplies (S_i) and demands (D_j) treated as bounded variables developed by Charnes and Klingman is extended to the case where the S_i and D_j are independently and uniformly distributed random variables. Chance constraints which require that demand at the jth destination will be satisfied with probability at least β_j and that stockout at the ith origin will occur with probability less than α_i are imposed. Conversion of the chance constraints to their linear equivalents results in a transportation problem with one more row and column than the original with some of the new arcs capacitated. The chance-constrained formulation is extended to the transshipment problem.

INTRODUCTION

Developments in network models have evolved from the first attempts by Dantzig [5] to solve transportation models using the simplex method through the stepping stone method of Charnes and Cooper [9] to recent advances in solution techniques for generalized networks (Balas [1], Balas and Hammer [2, 3, 4,], Charnes and Kirby [11], Charnes and Raike [12]). Considerable emphasis has been placed on developing efficient computational techniques. Lemke's [22] dual method, Dantzig's row-column-sum method (see [16]), Orden's [24] characterization of the transshipment problem as a transportation model, Wagner's [27] techniques for capacitated networks, Ford and Fulkerson's network algorithms ([17] and particularly the out-of-kilter technique [18]) and Vogel's approximation method [25] represent significant advances. Recent developments include investigations of efficient dual methods by Glover, Klingman and Napier [19], computation of efficient initial solutions (Glover, Klingman and Napier [28], Napier [23], Glover, Karney, Klingman, Napier [21]), and an improvement of the out-of-kilter method (Barr, Glover and Klingman [5]).

The great strides made in computational methods for network problems in the past twenty years, coupled with comparable advances in computer technology in the same period, have resulted in special purpose computer codes that can solve extremely large network problems in seconds. Napier [23] reports solution times in the neighborhood of 30 seconds on a CDC 6600 for networks of 100 nodes (dense) and 200 nodes (non-dense) with up to 10000 arcs. Also, more recently, Ross,

Klingman and Napier [26] have examined the effect on the computational efficiency of various problem dimensions in transportation problems such as the number of variables, rectangularity, density, number of constraints and the variance and skewness of its objective function.

It is reasonable to conclude that these special purpose algorithms can solve problems as large as will be encountered in actual networks. In an industrial application for, say, a monthly shipment schedule many minutes, if not hours, might otherwise be devoted to solving a very large problem.

STOCHASTIC DEMANDS AND SUPPLIES

Thus far, developments in network models have been in the main deterministic. Charnes and Kirby [11] present a number of chance-constrained formulations which might be applied to networks. Our aim here, however, is to present a special purpose technique for chance constraints applied to transportation or transshipment problems. Briefly, the model permits chance constraints on shipments such that the probability that all demands be met at a destination is constrained below; the probability of a stockout at an origin is bounded above.

We shall proceed from the standard transportation model to Charnes and Klingman's [13] treatment of supply and demand as bounded variables. From there, the extension to the chance-constrained model is quite natural and results in a tableau which is structurally identical to the bounded variables case.

THE CAPACITATED DISTRIBUTION MODEL

The chance-constrained distribution model is an adaptation of the Charnes and Klingman [13] modification of the distribution problem to encompass upper and lower bounds on the requirements.

Let x_{ij} be the amount flowing from node i to node j, c_{ij} the cost of that flow, S_i the total supply available at node i, and D_j the total demand at node j. Finally, let $I=\{i|i=1, \ldots, m\}$ and $J=\{j|j=1, \ldots, n\}$. The standard transportation problem is then:

Minimize
$$\sum_{i\epsilon I}\sum_{j\epsilon J} c_{ij}x_{ij}$$

Subject to:
$$\sum_{j\epsilon J} x_{ij}=S_i,\ i\epsilon I$$

(1)
$$\sum_{i\epsilon I} x_{ij}=D_j,\ j\epsilon J$$

$$x_{ij}\geq 0,\ i\epsilon I,\ j\epsilon J$$

Charnes and Klingman consider S_i and D_j to be bounded variables, thus allowing some flexibility in the distribution program. In effect this permits transportation costs to determine somewhat the requirements at the nodes. Designate S_j as the upper bound on D_j and $\underline{D}_j$ as the lower bound. Then $\underline{D}_j\geq D_j\leq \overline{D}_j$ and, similarly, $\underline{S}_i\leq S_i\leq \overline{S}_i$ for all $i\epsilon I$, $j\epsilon J$. Now, append I and J so that $I'=\{i|i=1, \ldots, m+1\}$ and $J'=\{j|j=1, \ldots, n+1\}$.

Consider a destination node where

(2a)
$$\sum_{i\epsilon I} x_{ij}=D_j$$

where $\underline{D}_j \leq D_j \leq \overline{D}_j$

Let $x_{m+1,\,j} = D_j - \underline{D}_j$. Equation (2a) becomes

$$\sum_{i \epsilon I} x_{ij} + x_{m+1,\,j} = \overline{D}_j. \tag{3}$$

and the bound conditions (2b) can be rearranged so that

$$D_j \leq \overline{D}_j \rightarrow 0 \leq \overline{D}_j - D_j = x_{m+1,\,j} \tag{4a}$$

$$\underline{D}_j \leq D_j \Rightarrow -\underline{D}_j \geq -D_j \rightarrow \overline{D}_j - \underline{D}_j \geq \overline{D}_j - D_j = x_{m+1,\,j} \tag{4b}$$

or, combining (4a) and (4b)

$$0 \leq x_{m+1,\,j} \leq \overline{D}_j - \underline{D}_j \tag{5}$$

An analagous derivation holds for $S_i \leq S_i \leq S_i$ which results in

$$0 \leq x_{i,\,n+1} \leq \overline{S}_j - \underline{S}_j \tag{6}$$

Thus the modified transportation problem is a capacitated distribution problem of the form: Minimize:

$$\sum_{i \epsilon I'} \sum_{j \epsilon J'} c_{ij} x_{ij}$$

Subject to:

$$\begin{aligned}
&\sum_{j \epsilon J'} x_{ij} = \overline{S}_i \qquad i \epsilon I \\
&\sum_{i \epsilon I'} x_{ij} = \overline{D}_j \qquad j \epsilon J \\
&\sum_{j \epsilon J'} x_{m+1,\,j} = \sum_{j \epsilon J} \overline{D}_j = \overline{S}_{m+1} \\
&\sum_{i \epsilon I'} x_{i,\,n+1} = \sum_{i \epsilon I} \overline{S}_i = \overline{D}_{n+1} \\
&0 \leq x_{m+1,\,j} \leq \overline{D}_j - D_j \qquad j \epsilon J \\
&0 \leq x_{i,\,n+1} \leq \overline{S}_i - S_i \qquad i \epsilon I \\
&x_{ij} \geq 0 \qquad i \epsilon I,\ j \epsilon J
\end{aligned} \tag{7}$$

Note that the conditions on the summations of the capacitated variables assure that

$$\sum_{i \epsilon I'} \overline{S}_i = \sum_{j \epsilon J'} \overline{D}_j$$

and also provide directly the actual total flow through the system by way of $x_{m+1,\,n+1}$.

CHANCE-CONSTRAINED ADAPTATION

Suppose, instead of treating the supplies and demands as bounded variables, we assume that the S_i, $i \epsilon I$ are uniformly and independently distributed random variables in the intervals $[S_j,\ \overline{S}_j]$

and the D_j, $j \epsilon J$ independent uniform deviates in the intervals $[\underline{D}_j, \bar{D}_j]$. We now insist that *all* demands at the j^{th} node be met with at least probability β_j, or

$$P\{\sum_{i \epsilon I} x_{ij} \geq D_j\} \geq \beta_j \tag{8}$$

Such chance constraints are readily converted to linear inequalities for uniform deviates (cf. Charnes Cooper and Symonds [10] and Charnes and Cooper [6]). Let x be uniformly distributed in $[a, b]$. Then

$$f(x) = \frac{1}{b-a}$$

$$F(x) = \int_a^x \frac{1}{b-a}\, dy = \frac{x-a}{b-a}$$

so that

$$F(x) = \frac{x-a}{b-a} \geq \beta$$

becomes

$$x \geq a + \beta(b-a)$$

Since (8) is of the form

$$F_j(\sum_{i \epsilon I} x_{ij}) \geq \beta_j$$

we can apply (9) to yield:

$$\sum_{i \epsilon I} x_{ij} \geq \underline{D}_j + \beta_j (\bar{D}_j - \underline{D}_j)$$

Let

$$\tilde{D}_j = \underline{D}_j + \beta_j (\bar{D}_j - \underline{D}_j)$$

So that (9) becomes

$$\sum_{i \epsilon I} x_{ij} \geq \tilde{D}_j \tag{10}$$

We can assume that no purpose is served by shipping more to node j than the maximum possible demand. Hence,

$$\tilde{D}_j \leq \sum_{i \epsilon I} x_{ij} \leq \bar{D}_j \tag{11}$$

The concern at the source nodes is that there not be a stockout, i.e., that S_i will satisfy the shipping requirements from the node.

Let α_i be the permitted stockout probability so that $1\text{-}\alpha_i$ is the desired probability that S_i exceed the programmed shipments from node i. The chance constraint can be expressed as

$$P\{S_i \geq \sum_{j \epsilon J} x_{ij}\} \geq 1 - \alpha_i$$

$$1 - F_i(\sum_{j \epsilon J} x_{ij}) \geq 1 - \alpha_i$$

$$F_i(\sum_{j \epsilon J} x_{ij}) \leq \alpha_i$$

$$\sum_{j \epsilon J} x_{ij} \leq \underline{S}_i + \alpha_i (\bar{S}_i - \underline{S}_i) \tag{12}$$

Let

$$S^*_i = \underline{S}_i + \alpha_i(\overline{S}_i - \underline{S}_i)$$

We assume that it is undesirable to ship less than $\underline{S}_i$, the minimum availability at node i. Thus,

$$\underline{S}_i \leq \sum_{j \epsilon J} x_{ij} \leq S^*_i \tag{13}$$

From (11) and (13) it is clear that the linear form of the chance constraints results in a distribution problem where

$$\sum_{i \epsilon J} x_{ij} \text{ and } \sum_{i \epsilon I} x_{ij}$$

are bounded variables. Hence we can, by substituting $\tilde{D}_j$ for $\underline{D}_j$ and S^*_j f $\overline{Sor}_j$ in (7), express the chance-constrained distribution problem as:

Minimize:

$$\sum_{i \epsilon I'} \sum_{j \epsilon J'} \bar{c}_{ij} x_{ij}$$

Subject to:

$$\sum_{j \epsilon J'} x_{ij} = S^*_i$$

$$\sum_{i \epsilon I'} x_{ij} = \overline{D}_j \tag{14}$$

$$\sum_{j \epsilon J'} x_{m+1,\, j} = \sum_{j \epsilon J} \overline{D}_j = S^*_{m+1}$$

where $\bar{c}_{ij} = Ec_{ij}$.

$$\sum_{i \epsilon I'} x_{i,\, n+1} = \sum_{i \epsilon I} S^*_i = \overline{D}_{m+1}$$

$$0 \leq x_{m+1,\, j} \leq \overline{D}_j - \tilde{D}_j \qquad j \epsilon J$$

$$0 \leq x_{i,\, n+1} \leq S^*_i - \underline{S}_i \qquad i \epsilon I$$

$$x_{ij} \geq 0 \qquad i \epsilon I,\ j \epsilon J$$

EXTENSION TO THE TRANSSHIPMENT PROBLEM

The conversion of the transshipment problem to a transportation model has been shown formally by Orden [25] so we shall proceed here by example. Our conversion differs slightly from Orden's in that he deals with net flows at transshipment nodes by altering either the demands or supplies by the net amount while, for our purposes, altering the demand or the supply for nodes at which, respectively, demands and supplies exist is more satisfactory.

Consider the network of Figure 1(a). Node 1 is a pure source, node 4 is a pure sink while 2 and 3 are transshipment nodes. A zero-cost feedback flow at the transshipment nodes can be introduced, as shown in Figure 1(b) without altering the problem. The network of Figure 1(b) can be represented as a distribution problem as shown in Figure 2(a).

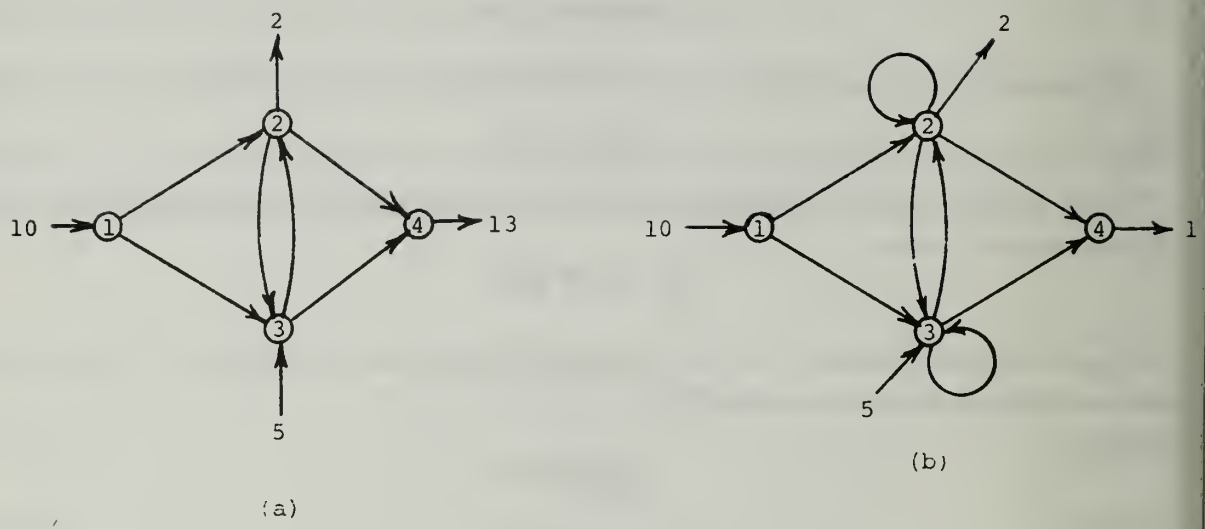


FIGURE 1.—A simple network.

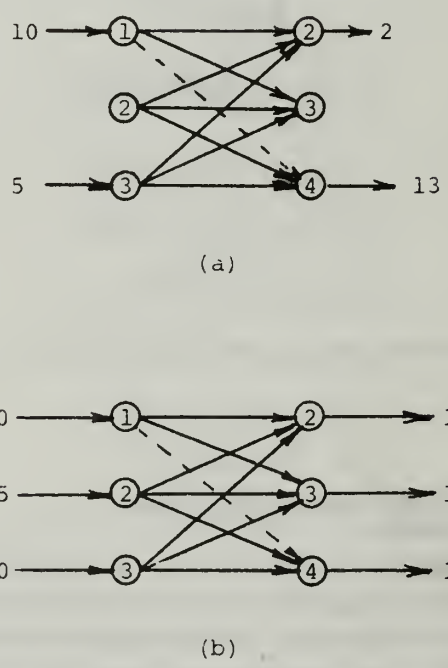


FIGURE 2.—Transportation model form of the simple network.

In the form of Figure 2(a), however it is possible that the feedback flows can take on negative values. This can be prevented by insisting that the total flows out of node 2 (at which a demand exists) equal the maximum of the possible inputs to node 2, *viz.* the sum of the total supplies to the system. Similarly the inputs to node 3 (at which there is supply) are equated to the total output from the system. The altered distribution problem is shown in Figure 2(b) where 15, the total flow, is added to both the inputs to and outputs from the transshipment nodes.

In algebraic form, the Kirchoff node conditions for the network of Figure 1(a) are shown in Table 1. Multiplying the equations for nodes at which demands exist by -1 we obtain the relations in Table 2.

We observe that the pure source (1) and the pure sink (4) are already in transportation problem form. We can replace the equations for the transshipment nodes (2) and (3) by pairs of equations whose difference yields the original equation, *viz.*

TABLE 1.—*Kirchoff Node Conditions for the Simple Network*

Arc → / Node	(1, 2)	(1, 3)	(1, 4)	(2, 2)	(2, 3)	(2, 4)	(3, 2)	(3, 3)	(3, 4)
1	$10=x_{12}$	$+x_{13}$	$+x_{14}$						
2	$-2=-x_{12}$				$+x_{23}$	$+x_{24}$	$-x_{32}$		$-x_{34}$
3	$5=$	$-x_{13}$			$-x_{23}$		$+x_{32}$		$+x_{34}$
4	$-13=$		$-x_{14}$			$-x_{24}$			$-x_{34}$

TABLE 2.—*Altered Kirchoff Node Conditions*

Arc → / Node	(1, 2)	(1, 3)	(1, 4)	(2, 2)	(2, 3)	(2, 4)	(3, 2)	(3, 3)	(3, 4)
1	$10=x_{12}$	$+x_{13}$	$+x_{14}$						
2	$2=x_{12}$				$-x_{23}$	$-x_{24}$	$+x_{32}$		$+x_{34}$
3	$5=$	$-x_{13}$			$-x_{23}$		$+x_{32}$		$+x_{34}$
4	$13=$		$+x_{14}$			$+x_{24}$			$+x_{34}$
	$15=$			x_{22}	$+x_{23}$	$+x_{24}$			
	$17=x_{12}$		$+x_{22}$			$+x_{32}$			
	and								
	$15=$	x_{13}			$+x_{23}$		$+x_{33}$		
	$20=$					$+x_{32}$	$+x_{33}$	$+x_{34}$	

In each case we have added an equation and a new variable representing the feedback node thus leaving the determination of the system unchanged. The problem constraints can be represented in tableau form as shown in Table 3. The non-existent Arc (1, 4) is shown with an arbitrarily large cost, M and the feedback nodes with 0 cost.

In brief, the conversion here first sets S_i equal to the supply at node i for pure sources and D_j equal to demand for pure sinks. Then

$$B=\sum_{i\epsilon I} S_i=\sum_{j\epsilon J} D_j$$

B computed and the demand at transshipment node j is set equal to $B+\min(0, D_j)$ and the supply equal to $B+\min(0, S_i)$.

TABLE 3.—*Simple Network in Transportation Model Form*

From—	To— 2	3	4	S_i
1			$\boxed{M}$	10
2	$\boxed{0}$			15
3		$\boxed{0}$		20
D_j	17	15	13	45

THE CHANCE-CONSTRAINED TRANSSHIPMENT MODEL

Let I_1 be the set of pure source nodes, J_3 the set of pure sinks and let I_2 and I_3 represent transshipment nodes which are sources and sinks, respectively, and J_1 and J_2 the destinations which are source and sink transshipment nodes. Clearly the problem:

Minimize:

$$E \sum_{i \epsilon I} \sum_{j \epsilon J} c_{ij} x_{ij}$$

Subject to:

$$\sum_{j \epsilon J} x_{ij} = S_i \qquad i \epsilon I_1$$

$$\sum_{j \epsilon J} x_{ij} = B + S_i \qquad i \epsilon I_2$$

$$\sum_{j \epsilon J} x_{ij} = B \qquad i \epsilon I_3$$

(15)

$$\sum_{i \epsilon I} x_{ij} = B \qquad j \epsilon J_1$$

$$\sum_{i \epsilon I} x_{ij} = B + D_j \qquad j \epsilon J_2$$

$$\sum_{i \epsilon I} x_{ij} = D_j \qquad j \epsilon J_3$$

$$x_{ij} \geq 0 \qquad i \epsilon I,\ j \epsilon J$$

The chance-constrained problem, by substituting $B+S_i$ or $B+D_j$ for S_i and D_j in (8) and (12) as required is readily comprehended as

Minimize:

$$\sum_{i \epsilon I'} \sum_{j \epsilon J'} \bar{c}_{ij} x_{ij}$$

Subject to:

$$\sum_{j \epsilon J'} x_{ij} = S^*_i \qquad i \epsilon I_1$$

$$\sum_{j\epsilon J'} x_{ij}=B+S^*_i \qquad i\epsilon I_2$$

$$\sum_{j\epsilon J'} x_{ij}=B \qquad i\epsilon I_3$$

$$\sum_{i\epsilon I'} x_{ij}=B \qquad j\epsilon J_1$$

$$\sum_{i\epsilon I'} x_{ij}=B+\overline{D}_j \qquad j\epsilon J_2$$

$$\sum_{i\epsilon I'} x_{ij}=\overline{D}_j \qquad j\epsilon J_3$$

(16)

$$\sum_{j\epsilon J'} x_{m+1,\,j}=\sum_{j\epsilon J} \overline{D}_j=S^*_{m+1}$$

$$\sum_{i\epsilon I'} x_{i,\,n+1}=\sum_{i\epsilon I} S^*_1=\overline{D}_{m+1}$$

$$0\leq x_{m+1,\,j} \qquad j\epsilon J_1$$

$$0\leq x_{m+1,\,j}\leq \overline{D}_j-\tilde{D}_j \qquad j\epsilon J_2 U J_3$$

$$0\leq x_{i,\,n+1}\leq S^*_1-\underline{S}_i \qquad i\epsilon I_1 U I_2$$

$$0\leq x_{i,\,n+1} \qquad i\epsilon I_3$$

$$x_{ij}\geq 0 \qquad i\epsilon I,\ j\epsilon J$$

where, as usual, it is assumed that c_{ij} and x_{ij} are independent so that $Ec_{ij}x_{ij}=c_{ij}x_{ij}$.

CONCLUSIONS

The addition of chance constraints to the transportation and transshipment models has been accomplished with only a trivial increase in computational requirements. In an $m\times n$ transportation model the conversion involves the addition of $m+n+1$ arcs where at most max (m, n) of these are capacitated. Thus, large scale chance-constrained distribution models can be solved using already available computer codes.

RECOMMENDATIONS FOR FURTHER RESEARCH

The chance constraints investigated here with uniform random deviates can be adapted readily to any density of the S_i and D_j provided, of course: (1) the variables are independently distributed; (2) the density functions are non-zero only in a finite interval (i.e., truncated above and below); and (3) the inverse cumulative distribution function can be computed or derived by Monte Carlo methods. The triangular and beta densities immediately present themselves as do empirically derived densities which are inherently truncated.

In its present form the model belongs to the class of zero order decision rule chance-constrained models. Thus, further work must be done to expand the model in a dynamic context to maximizing over a multi-period horizon.

BIBLIOGRAPHY

[1] Balas, Egon. "The Dual Method for the Generalized Transportation Problem," Management Science, *12*, 555–568 (1966).

[2] Balas, Egon and P. L. Hammer (Ivanescu), "On the Generalized Transportation Problem," Management Science, *11*, 188–202 (1964).

[3] Balas, Egon and P. L. Hammer (Ivanescu), "On the Transportation Problem—Part I," Cahiers du Centre d'Etudes de Recherche Operationelle, *4*, No. 2 (1962).

[4] Balas, Egon, and P. L. Hammer (Ivanescu), "On the Transportation Problem—Part II," Cahiers du Centre d'Etudes de Recherche Operationelle, *4*, No. 3 (1962).

[5] Barr, R. S., F. Glover and D. Klingman, "An Improved Version of the Out-of-Kilter Method and A Comparative Study of Computer Codes," Mathematical Programming, *7*, 60–86 (1974).

[6] Charnes, A. and W. W. Cooper, "Chance-Constrained Programming," Management Science, *6*, 73–79 (1959).

[7] Charnes, A. and W. W. Cooper, "Deterministic Equivalents for Optimizing and Satisficing under Chance Constraints," Operations Research, *11*, 18–39 (1963).

[8] Charnes, A. and W. W. Cooper, *Management Models and Industrial Applications of Linear Programming*. (New York: John Wiley & Sons, Inc., 1961).

[9] Charnes, A. and W. W. Cooper, "The Stepping Stone Method of Explaining Linear Programming in Transportation Problems," Management Science, *1*, No. 1 (1954).

[10] Charnes, A., W. W. Cooper, and G. H. Symonds, "Cost Horizons and Certainty Equivalents: An Approach to Stochastic Programming of Heating Oil," Management Science, *4*, 235–263 (1958).

[11] Charnes, A. and M. Kirby, "The Dual Method and the Method of Balas and Ivanescu for the Transportation Model," Cahiers du Centre d'Etudes de Recherche Operationelle, *6*, No. 1 (1964).

[12] Charnes, A. and M. J. L. Kirby, "Some Special P-Models in Chance-Constrained Programming," Management Science, *14*, 183–195 (1967).

[13] Charnes, A and D. Klingman, "The Distribution Problem with Upper and Lower Bounds on Node Requirements," Management Science, *16*, 638–642 (1970).

[14] Charnes, A. and W. M. Raike, "One-Pass Algorithm for Some Generalized Network Problems," Operations Research, *14*, 914–924 (1966).

[15] Dantzig, G. B., "Application of the Simplex Method to a Transportation Problem," in T. C. Koopmanns (ed.), *Activity Analysis of Production and Allocation*, Cowles Commission Monograph No. 13. (New York: John Wiley & Sons, Inc., 1951).

[16] Dantzig, G. B., *Linear Programming and Extensions* (Princeton, N.J.: Princeton University Press, 1963).

[17] Ford, L. R., Jr., and D. R. Fulkerson, *Flows in Networks* (Princeton, N.J.: Princeton University Press, 1962).

[18] Ford, L. R., Jr., and D. R. Fulkerson, "An Out-of-Kilter Method for Minimal Cost Flow Problems," SIAM Journal, *9*, No. 1 (1961).

[19] Glover, F., D. Klingman and A. Napier, "An Efficient Dual Approach to Network Problems," Working Paper 71–57, The University of Texas, Austin, Texas (May 1971).

[20] Glover, F., D. Klingman and A. Napier, "A One-Pass Algorithm to Determine a Dual Feasible Basic Solution for a Class of Capacitated Generalized Networks," Center for Cybernetic Studies, Research Report 42, The University of Texas, Austin, Texas (October, 1970).

[21] Glover, F., D. Karney, D. Klingman and A. Napier, "A Computation Study on Start Procedures, Basis Change Criteria, and Solution Algorithms for Transportation Problems," Management Science, *20*, 793–813 (1974).

[22] Lemke, C., "The Dual Method of Solving Linear Programming Problems," Naval Research Logistics Quarterly, *1*, No. 1 (1954).

[23] Napier, H. A. Jr., "Some Algorithmic Procedures for Networks and their Computational Relationship with Existing Network Algorithms." Doctoral Dissertation, The University of Texas, Austin, Texas (May 1971).

[24] Orden, A., "The Transshipment Problem," Management Science, *2*, 276–285 (1956).

[25] Reinfield, N. V. and W. R. Vogel, Mathematical Programming (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1958).

[26] Ross, Terry G., D. Klingman and A. Napier, "A Computational Study of the Effects of Problem Dimensions on Solution Times for Transportation Problems," Journal of the Association for Computing Machinery, *22*, 413–424 (1975).

[27] Wagner, H. M., "On a Class of Capacitated Transportation Problems," Management Science, *5*, 304–318 (1959).

[28] Wagner, H. M., *Principles of Operations Research.* (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1969).

ELEMENTS OF A THEORY IN NON-CONVEX PROGRAMMING

Claude-Alain Burdet

SYSTEMATHICA Consulting Group Ltd.
Pittsburgh, Pennsylvania

ABSTRACT

The question of necessary and sufficient optimality conditions for non-convex programs is analyzed in the general context of subadditivity. Several types of convex set extensions are investigated to generate valid inequalities from the corresponding gauge functions.

1. SUMMARY

We first present a set of "naive" *optimality conditions* (necessary and sufficient) applicable to a general mathematical program. An example and its Kuhn-Tucker conditions are described to show why such necessary conditions are impractical in a non-convex situation.

We then proceed with necessary conditions in inequality form and show how this concept is related to that of *subadditive gauge* functions (see Section 3.2). In Sections 3.3 and 3.4, we generalize Tuy's intersection method and describe the construction of cutting planes primarily based on the objective function; it is shown that Tuy's cuts are uniformly dominated in the present framework. Dominance is strict when the level set of the objective function is unbounded.

We next investigate Tuy's idea of convex extension. A new formulation allows relaxation of the convexity assumption for the objective function and produces *quasi-convex extensions* which subsume Tuy's concept. In fact extensions can be defined in very general terms (see Section 3.6); and any valid inequality (including the most stringent ones) can be cast in this framework.

Section 4 investigates an analytical characterization of the concept of extension and the possibility to explicitly construct the corresponding cutting planes. This is obtained from convex, quasi-convex and/or polaroid gauge extensions. The basic thrust of our research is to obtain results stronger than Tuy's cuts which have proved disappointing. We use subadditive gauges on the one hand and extensions on the other as a vehicle to circumvent the obstacles met by Tuy's concepts and methodology. Although no algorithmic implication is discussed here in depth, the general direction of this study is aimed at solving some typical problems in the difficult area of non-convex optimization.

2. OPTIMALITY

Consider the mathematical program
Maximize $f(x)$ subject to $x \epsilon X \subset R^n$

One has the immediate *necessary and sufficient optimality conditions:*

$$\left.\begin{array}{l}\bar{x}\epsilon X \text{ is optimal} \\ \text{with } \bar{k}=f(\bar{x})\end{array}\right\} \text{iff } X\subset \operatorname{lev}_{\bar{k}} f \dagger \tag{2.1}$$

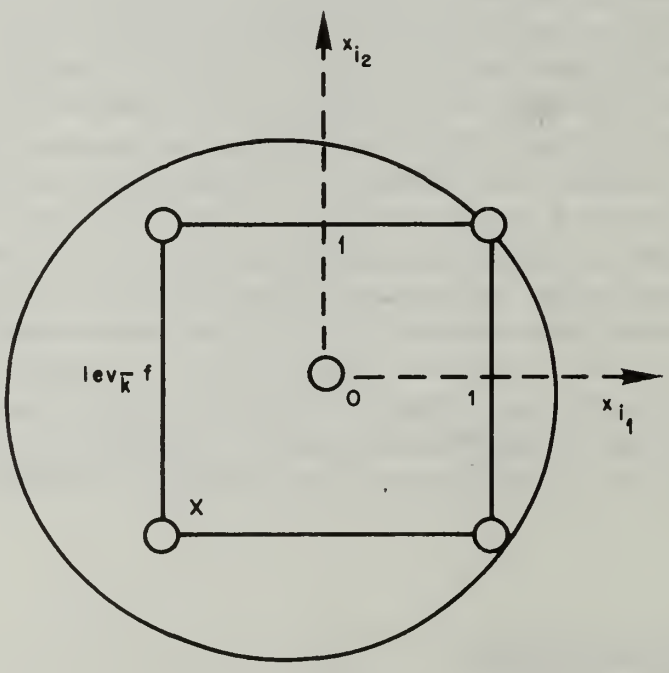
The above optimality conditions are "naive" in the sense that they yield no constructive solution to the mathematical program and, in fact, they represent little more than a parody of the definition for *global optimality.*

It is well known in the non-linear programming literature that *necessary conditions for local optimality* can be obtained from the Kuhn-Tucker theory; furthermore, under suitable convexity assumptions for f and X, these conditions turn out to be sufficient, so that the question of optimality is fully answered.

In the absence of convexity, the situation is quite different however: sufficiency is no longer guaranteed. Moreover there frequently exists an inordinate number of local optima so that explicit search is impractical. But the situation is worse yet (as shown in the illustrative example (Figure 1a) below): in addition to numerous local optima, there also exists a myriad of "useless" Kuhn-Tucker points of the saddle type, which are not locally optimal! Thus global optimization by means of successive inspection of K-T points tends to be very inefficient.

EXAMPLE: Consider the "concave programming" problem:

Maximize $f=x^2+\epsilon x$
subject to $-1\leq x_i\leq 1, \quad i=1, \ldots, n$
where ϵ_i are small positive quantities $\forall i$.



Optimal value:

$$n+\sum_i \epsilon_i=\bar{k}$$

Optimal Solution:

$$x_i=1, \forall_i$$

Figure 1a

†The notations lev, aff, cl, epi, bd and conv denote the level set, affine hull, closure, epigraph, boundary set and convex hull respectively [8].

Number of Kuhn-Tucker points: $\sum_{i=0}^{n} \binom{n}{i} 2^{n-i}=9$

Number of local minima (global): $\binom{2}{2} 2^0=1$ (convex case)

Number of local maxima: $\binom{2}{2} 2^2=4$

Number of saddle points: $\sum_{i=1}^{n-1} \binom{n}{i} 2^{n-i}=\binom{2}{1} 2=4$

In this case there is a Kuhn-Tucker point on each face (of every dimension $k \leq n$) of the hypercube X; thus the Kuhn-Tucker theory requires one to check an enormous number of points, only very few of which are locally optimal; but even local optima are too numerous (2^n) to be checked explicitly as soon as n becomes reasonably large (say >20).

3. INEQUALITIES

3.1. Generalities

In the search for alternatives to the Kuhn-Tucker theory, one may attempt verifying (constructively) the inclusion $X \subset \text{lev}_{\bar{k}} f$.

This can be conceptually accomplished in many different ways which eventually amount to "splitting" X into several subsets with subsequent verification of the inclusion for each subset. Cutting planes are usually associated with a repeated "shrinking process" of the feasible set, which may lead to convergence difficulties; alternately another approach is described below: it is finite and therefore remains free of convergence problems. In practice a large number of subsets is usually required; but the use of cutting planes (i.e., necessary conditions) may in certain cases furnish some hope to efficiently curtail the search for global optimality.

EXAMPLE: Let X be defined by the inequalities

$$g_k(x) \leq 0, \; k \in M \tag{3.1}$$

The Kuhn-Tucker conditions are known to contain a set of complementarity conditions which read:

$$\lambda_k g_k = 0 \qquad \forall k \in M \tag{3.2}$$

This structure can now be used to set up a dichotomous arborescence which generates $2^{m+1}-1$ subproblems of X according to the two possibilities:

$$\text{a) } \lambda_k = 0, \; g_k < 0 \qquad \text{b) } g_k = 0, \; k \in M \tag{3.3}$$

which both imply (3.2) and are mutually exclusive.

This "facial" decomposition methodology [5] has been applied successfully to some classes of problems in integer programming and in general quadratic programming; it is particularly useful when X is polyhedral (i.e., g_k linear). Facial decomposition is a natural extension of the idea of Branch and Bound, as both methods become identical when X is a hypercube.

The use of cutting planes offers a possibility to discard certain subproblems from further explicit computations.

Since the proposed method is similar to Branch and Bound algorithms, it possesses the same underlying structure of exhaustive search among candidates for the global optimum. Thus the process is *intrinsically finite*. For a polyhedral feasible set X, one obtains a tree structure which is completely determined by the combinatorial structure of X; each node is a program on a face of X (subprogram). Nodes are fathomed by a machinery of cutting planes designed to eliminate subprograms (faces) known to contain no point with a better objective function value than the current best solution (see [5] for more details).

3.2. Inequalities Generated From Subadditive Gauge Functions

We now investigate a general framework for the construction of inequalities.

DEFINITION 1: A *subadditive gauge** function π is defined to satisfy:

SUBADDITIVITY:

$$\pi(u)+\pi(v)\geq\pi(u+v),\ \forall u,\ v \tag{3.4a}$$

GAUGE:

$$\pi(\lambda u)=\lambda\pi(u),\ \forall\lambda\geq 0 \text{ (positively homogeneous of degree one)} \tag{3.4b}$$

LEMMA 1: For gauge functions, subadditivity and convexity are equivalent.

PROOF:

From convexity:

$$\pi(\alpha u+(1-\alpha)v)\leq\alpha\pi(u)+(1-\alpha)\pi(v)$$

hence for

$$\alpha=\frac{1}{2}$$

one obtains

$$\pi\left(\frac{1}{2}(u+v)\right)=\frac{1}{2}\pi(u+v)\leq\frac{1}{2}(\pi(u)+\pi(v)),$$

i.e. subadditivity

Conversely:

$$\alpha\pi(u)+(1-\alpha)\pi(v)=\pi(\alpha u)+\pi((1-\alpha)v)\geq\pi(\alpha u+(1-\alpha)v) \qquad \text{Q.E.D.}$$

For simplicity we now assume that every point $x\in X$ can be represented by reference to a given finite system of vectors $\{e_j\in R^n\}$ with respect to $\tilde{x}\in$ Aff $X\subset R^n$ i.e.

$$x=\tilde{x}+u=\tilde{x}+\sum_j t_j e_j.$$

Assume the "coordinates" t_j to be *non-negative*, i.e. the system $\{e_i\}$ spans a cone at $\tilde{x}$ which contains X.

---

*Rockafellar [8] introduced this gauge function terminology in the context of Minkowski functionals characterizing a convex body; the present concept is closely related and we use the same term to characterize a broader class of functions.

DEFINITION 2: An inequality

$$\sum_{j\in N} \pi_j t_j \geq \pi_o$$

is called *k-valid* if the set

$$X \cap \left\{ x=x(t) \middle| \sum_{j\in N} \pi_j t_j < \pi_o \right\} \tag{3.5}$$

contains no point x such that $f(x) \geq k$.

THEOREM 1: Let π be a subadditive gauge: $\{\text{Aff } (X) - \tilde{x}\} \to R$; then the inequality

$$\sum_{j\in N} \pi_j t_j \geq \pi_o$$

is k-valid,

$$\text{with } \pi_j = \pi(e_j), \tag{3.6a}$$

$$\pi_o = \min_{\substack{x\in X \\ f(x)\geq k}} \pi(u) \text{ where } u = \sum e_j t_j = x - \tilde{x}, \ \forall x \in X \tag{3.6b}$$

PROOF: By induction on n. Since π is a gauge and $t_j \geqq 0$, $\forall j \in N$, one has:

$$\pi_{j_1} t_{j_1} = \pi(e_{j_1}) t_{j_1} = \pi(e_{j_1} t_{j_1})$$

$$\begin{array}{ccccccc}
\pi(e_{j_1}t_{j_1}) & + & t_{j_2}\pi_{j_2} & = & \pi(e_{j_1}t_{j_1}) + \pi(e_{j_2}t_{j_2}) & \geq & \pi(e_{j_1}t_{j_1} + e_{j_2}t_{j_2}) \\
 & & \vdots & & \vdots & & \\
\cdots & \cdots & = & \cdots & \cdots & \geq & \pi\left(\sum_j e_j t_j\right) = \pi(u)
\end{array}$$

Summing all inequalities, one obtains:

$$\sum_j \pi_j t_j \geq \pi(u).$$

Now $\pi(u) \geq \pi_o$, $\forall x \in X$ with $f(x) \geq k$, by definition. Q.E.D.

REMARK: Since one usually does not know in advance the optimal value $\bar{k}$, one has to construct k-valid inequalities with $k \leqq \bar{k}$, where k is a known (or estimated) *lower bound* for the maximum $\bar{k}$.

EXAMPLE: Let X be characterized by the linear system

$$x_i = \sum_j a_{ij} t_j \leq \tilde{x}_i, \ i \in M$$

$$t_j \geqq 0$$

as customary in linear programming

Then one may define as in (3.6a):

$$\pi(e_j) = \pi(a_j), \text{ with } e_j = a_j = (a_{1j}, \ldots, a_{ij}, \ldots, a_{mj})$$

$$\pi_o = \min_{\substack{u=x-\tilde{x} \\ f(x)\geq k \\ x\in X}} \pi(u)$$

For instance, choose (see [3])

$$\pi=\sum_{i\in M}\rho_i(u_i)|u_i|$$

with parameters

$$\rho_i(u_i)=\begin{cases}\rho_i^+, \text{ if } u_i\geq 0\\ -\rho_i^-, \text{ if } u_i<0\end{cases}\qquad \forall i\in M$$

where $\rho_i^+\geq\rho_i^-$ for convexity of π (hence subadditity since π is clearly a gauge) i.e.

(3.7a) $$\pi_i=\sum_{i\in M}\rho_i(a_{ij})|a_{ij}|$$

(3.7b) $$\pi_o=\min_{\substack{u=x-\tilde{x}\\ f(x)\geq k}}\sum_{i\in M}\rho_i(u_i)|u_i|$$

This introductory example exhibits several features which are characteristic of the approach developed here:

a) The subadditive gauge π is an auxiliary tool which can be chosen independently of the problem (i.e. f and X). Of course there may exist some practical advantages in constructing "tailor made" gauges which reflect the structure of f and/or X. Indeed the program (3.6b) may be very difficult for arbitrary functions π. Most of the remaining developments of this paper are dealing with the question of exploiting any information contained in f and/or X to construct "appropriate" gauges in that respect.

b) The coefficients π_j of the inequality may be >0, $=0$ or <0; thus the theory is able to generate any type of valid inequality.

c) The amount of computations required for the construction of the inequality lies in the determination of π_j and π_o; the above example shows, however, that the computations for π_j are of a different nature than for π_o. Since there are n coefficients π_j, one will typically choose π functions where π_j is easily obtained (as in the example); the major part of the computational effort then consists in the determination of the coefficient π_o. Note that a lower bound $\pi_o\leq\overline{\pi}_o$ is sufficient to yield a k-valid inequality.

3.3. Subadditive Gauges Directly Related to f

Consider the "concave programming" problem

Maximize $f(x)$

Subject to $x\in X\subset R^n$

where f is quasi-*convex*. Assume that a value k ($\leq\bar{k}$) and a point $\tilde{x}\in$ Aff X with $f(\tilde{x})\leq k$ are known. We now construct a gauge function π based on the convex set (lev $_k f)-\tilde{x}$: (see Figure 1b)

(3.8a) a) π is a convex gauge, in particular $\pi(0)=0$

(3.8b) b) $\forall x\neq\tilde{x}$ such that $f(x)=k$, let $u=x-\tilde{x}$,

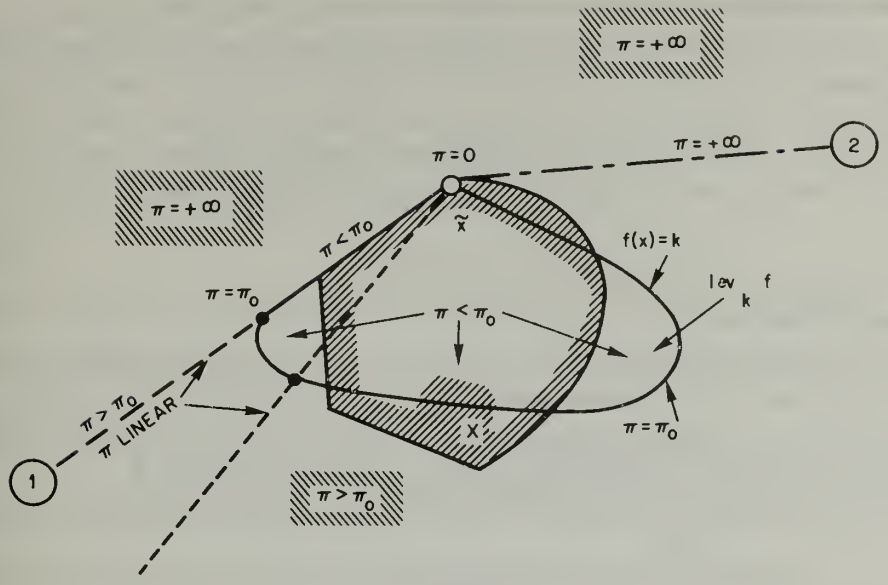


FIGURE 1b.—Construction of an f related gauge (see (3.8)). The case (3.8a) applies to the point $\tilde{x}$. The case (3.8b) applies to the top half of the illustration, including the line 2 which is tangential to the set X at $\tilde{x}$. The case (3.8c) applies to the lower half, including the line 1. One has $\pi<\pi_o$ within the entire set $\text{lev}_k f$; along the boundary of $\text{lev}_k f$, where $f(x)=k$, one has $\pi=\pi_o$ (curved part) and $\pi\leq\pi_o$ (linear part).

and set

$$\text{(3.8c)}\qquad \pi(u)=\begin{cases}+\infty, \text{ if } f(\lambda x+(1-\lambda)\tilde{x})=f(\tilde{x}+\lambda u)>k, \ \forall\lambda>0\\ \pi_o/\lambda^* \text{ otherwise, where } \lambda^*=\max\limits_{0f\lambda<+\infty}\{f(\lambda x+(1-\lambda)\tilde{x})\leq k\}, \text{ i.e. } \lambda^*\geq 1.\end{cases}$$

Note that when $f(\tilde{x})<k$, and $\text{lev}_k f$ is a bounded set the condition b) reduces to:

$$\pi(u)=\pi_o, \ \forall x \text{ such that } f(x)=k;$$

indeed convexity of $\text{lev}_k f$ implies that $\lambda=1$ is the unique solution of

$$f(\lambda x+(1-\lambda)\tilde{x})=k.$$

Choose a system $\{e_j\}$ at $\tilde{x}$ such that

$$\forall x\in X: x=\tilde{x}+\sum_j e_j t_j, \ t_j\geq 0$$

we now have:

LEMMA 2:

$$\text{(3.8d)}\qquad \pi(u)\leq\pi_o, \ \forall u\in[(\text{lev}_k f)-\tilde{x}] \text{ i.e. } \text{lev}_{\pi_o}\pi=[(\text{lev}_k f)-\tilde{x}]$$

PROOF: Take x with $f(x)>k$. Since π is a gauge, one has on the ray λu, with $u=x-\tilde{x}$ and $\lambda\in[0, 1]$:

$$\pi(\lambda u)=\lambda\pi(u).$$

But by hypothesis $f(\tilde{x}) \leq k$; hence for some $\lambda \epsilon [0, 1)$, one has $f(\lambda u + \tilde{x}) = k$:

a) if $\lambda = 0$ is the only such value, then $\pi(\lambda u) = \pi(u) = +\infty > \pi_o$, $\forall \lambda > 0$ (from 3.8b); and $\tilde{x} \epsilon bd \ \mathrm{lev}_{\pi_o} \pi$.

b) Otherwise for all λ with $f(\lambda u + \tilde{x}) = k$, one has $\lambda \pi(u) \leq \lambda^* \pi(u) = \pi_o$, from (3.8c);

furthermore (3.8c), i.e. $\lambda^* \leq 1$ and $f(x) < k$, imply $\lambda^* < 1$. Hence one must have $\pi(u) > \pi$. Q.E.D.

LEMMA 3: Assume local optimality of $\tilde{x}$ i.e. $f(\tilde{x} + \lambda e_j) \leq f(\tilde{x})$, $\forall \lambda \epsilon [0, \epsilon)$ with $\epsilon > 0$ sufficiently small. Then

a) $\pi_j = \pi_o / \lambda_j^*$, with $\lambda_j^* = \max_{0 \leq \lambda < +\infty} \{ f(\lambda e_j + \tilde{x}) = k \}$

b) $\pi_j \leq 0$ otherwise

PROOF: a) by construction one has $\lambda_j^* \pi_j = \pi(e_j \lambda_j^*) = \pi_o$; since local optimality of $\tilde{x}$ implies that for some $\lambda > 0$, the case $\pi = +\infty$ cannot occur.

b) If $f(e_j \lambda_j + \tilde{x}) < k$, $\forall \lambda_j > 0$, one may always set $\pi_j = 0$, i.e. $\lambda_j^* = \infty$; however non-negativity of the π gauge is not assumed, nor implied by the convexity assumption (3.4). Q.E.D.

THEOREM 2: Assume (3.8) to hold true and let $\tilde{x}$ be locally optimal with $f(\tilde{x}) \leq k$. Then

$$\sum_{j \epsilon N} \pi_j t_j \geq \pi_o$$

is a k-valid inequality, where

(3.9a) a) $\pi_j = \pi_o / \lambda_j^*$, $\lambda_j^* = \max_{0 < \lambda \leq +\infty} \{ f(\tilde{x} + \lambda e_j) = k \}$

(3.9b) b) $\pi_j \leq 0$, for the other $j \epsilon N$ where (3.9a) does not apply.

PROOF: Follows from the Lemmas 2 and 3. Q.E.D.

REMARK: Theorem 2 remains true if $\tilde{x}$ is only assumed to satisfy $f(\tilde{x}) < k$.

COROLLARY 2.1 (Tuy): Assume $f(\tilde{x}) < k$. The cut

$$\sum_{j \epsilon N} \hat{\pi}_j t_j \geq \hat{\pi}_o$$

is k-valid, with

(3.10) $$0 \leq \pi_j = \pi_o / \hat{\lambda}_j,$$

where

$$\hat{\lambda} = \max_{0 < \lambda \leq +\infty} \{ f(\tilde{x} + \lambda e_j) \leq k \} \leq +\infty.$$

PROOF: We need verify the hypothesis (3.8) of Theorem 2: (3.8b) is immediate from Lemma 3 which contains the construction (3.10) as a special case.

We now show that the resulting gauge π is *convex* and therefore satisfies (3.8a). π is constructed with convex level sets (3.8d) and is therefore quasi-convex; it is also non-negative. Lemma 4 below completes the proof.

Q.E.D.

LEMMA 4: A non-negative quasi-convex gauge π is convex.

PROOF: We show that the set epi $\pi = \{ (\pi, u) | \pi \geq \pi(u), u \epsilon R^n \}$ is a *convex* cone in R^{n+1}. Since π is a *gauge*, one has $\forall u \epsilon R^n$ with $0 < \pi(u) < +\infty : \pi(u) = \pi(\mu v) = \mu \pi(v) = \mu$
where $u = \mu v$, and $\mu > 0$ with $v \epsilon$ bd $\mathrm{lev}_1 \pi$, i.e. $\pi(v) = 1$
Thus all (strictly) positive values of π are completely determined by the set $\mathrm{lev}_1 \pi$.

Since $\text{lev}_1\,\pi$ is *convex*, by the quasi-convexity assumption on π, there exists (at least) one hyperplane $H_v u=1$ which supports $\text{lev}_1\,\pi$ at the point v, i.e. $H_v u \geq 1$, $\forall u \in \text{lev}_1\,\pi$ and $H_v v=1$.
Since the gauge π is *non-negative* the set epi π can be represented as an intersection of the following convex sets.

$$\text{epi}\,\pi = \bigcap_{v \in \text{bd lev}_1\,\pi} \{(h, u) | h \geq \max\{0, H_v u\}, u \in R^n\}$$

Q.E.D.

3.4. Dominance

As indicated by the above Corollary, one sees that Tuy's theory is generalized in the present framework:

(i) For a given level set $\text{lev}_k f$, one constructs in a straightforward manner the subadditive gauge which corresponds to Tuy's cut (see Corollary 2.1).

(ii) Further k-valid inequalities can be obtained, however, when $\text{lev}_k f$ contains unbounded rays $(t_j e_j + \tilde{x})$, $j \in N$: Any convex gauge π which agrees identically with $\hat{\pi}$ on the set R_I of *intersected directions*
$R_I = \{u \in [\text{Aff } X - \tilde{x}] | f(\lambda u + \tilde{x}) > k$, for some $\lambda \geq 1\}$ will yield such an inequality.

EXAMPLE (See Figure 2): $f=(xy)^{-1}$, $k=1/2$; $\pi_o=1$, $\tilde{x}=(2, 1)$ i.e. $x=u+2$, $y=v+1$.

A) $u \leq 0$ or $v \leq 0$: in the construction (3.8b) we must find $\hat{\lambda}$ from $[(\hat{\lambda}u+2)(\hat{\lambda}v+1)]^{-1}=1/2$
i.e., $\hat{\lambda}^2 uv + \hat{\lambda}(u+2v)=0$ and thus

$$\hat{\lambda} = \frac{-(u+2v)}{uv}.$$

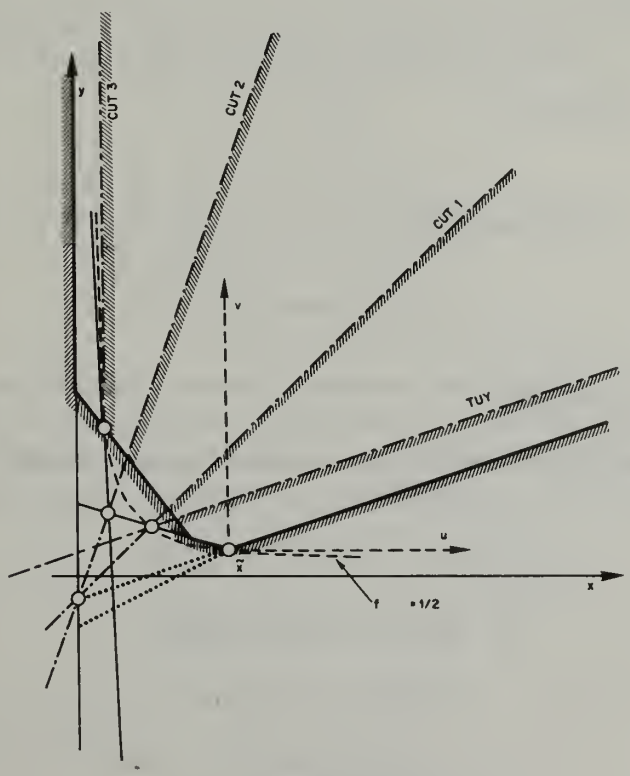


FIGURE 2a

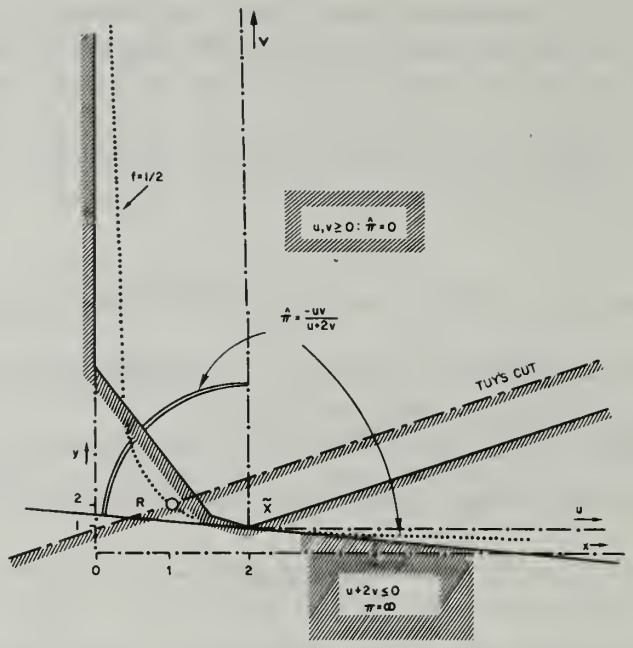


FIGURE 2b.—Illustration of the construction of Tuy's cut. One intersection point is R, the other at infinity, because the gauge $\hat{\pi}$ vanishes in $u, v, >0$. The point R can be determined geometrically by intersecting the ray r with the curve.

$$f=\frac{1}{2}.$$

Now in the region

$$\frac{1}{2}u+v\geq 0$$

one has $\pi=\hat{\lambda}^{-1}$; $\pi=+\infty$ otherwise.

Thus:

$$\pi(u, v)=\frac{-uv}{(u+2v)}, \begin{cases}\text{for } u\geq 0,\ v\leq 0 \text{ and } \frac{1}{2}u+v\geq 0\\ \text{for } u\leq 0,\ v\geq 0,\ \frac{1}{2}u+v\geq 0\end{cases}$$

$$\pi(u, v)=+\infty \qquad \text{for } \frac{1}{2}u+v<0$$

$$\pi(0, 0)=0$$

B) $u\geq 0, v\geq 0$ (Tuy): $\hat{\pi}(u, v)=0$. But other gauges can be obtained as follows:

$$\frac{\partial\pi}{\partial u}=\frac{-v}{u+2v}+\frac{uv}{(u+2v)^2}=\frac{-v^2}{(u+2v)^2}$$

$$\frac{\partial\pi}{\partial v}=\frac{-u}{(u+2v)}+\frac{2uv}{(u+2v)^2}=\frac{-u^2}{(u+2v)^2};$$

for

$$u=0: \frac{\partial\pi}{\partial u}=\frac{-v^2}{2v^2}=-\frac{1}{2}, \forall v;$$

and for

$$v=0: \frac{\partial\pi}{\partial v}=\frac{-u^2}{u^2}=-1, \forall u.$$

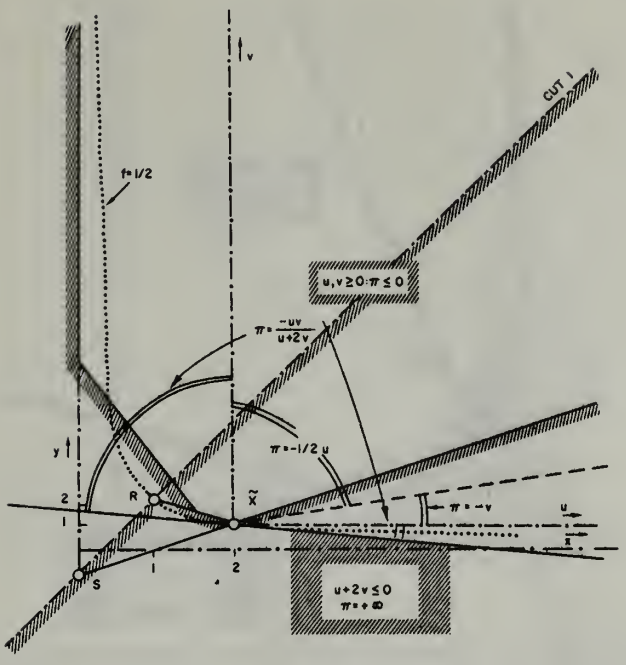


FIGURE 2c.—Cut 1 differs from Tuy's cut in that it now passes through the points S and R; in Tuy's cut the point S was at ∞, while here the function.

$$\pi(u, v) = -\frac{1}{2}u \leq 0$$

defines S. The polaroid gauge construction (4.8) corresponds geometrically to the envelope, within the non-negative quadrant $u, v \geq 0$, of all tangent hyperplanes to the set epi f along the u and v axes. In the remaining region of R^2 with $u+2v \geq 0$, the polaroid gauge π remains identical to

$$\hat{\pi} = \frac{-uv}{u+2v}.$$

Thus, in the non-negative quadrant, one can choose any value satisfying:

$$0 \geq \pi(u, v) \geq \max\left\{-\frac{1}{2}u, -v\right\}.$$

The analytical derivation of this construction is based on *polaroid extensions* (See Section 4.2).

One easily verifies that the functions π defined by A) and B) are all convex gauges with cl $(\text{lev}_1 \pi) = (\text{lev}_1 f - \tilde{x})$. Note that there are, in general, many possibilities to define suitable gauges in the region $u \geq 0$, $v \geq 0$; all generalize Tuy's cut; therefore it is natural to investigate the "best" such generalization.

THEOREM 3: (dominance) Assume $\pi_o > 0$. The k-valid inequality

$$\sum_j \pi_j t_j \geq \pi_o, \text{ where } t_j \geq 0$$

dominates uniformly any inequality

$$\sum_j \alpha_j t_j \geq \pi_o$$

iff $\alpha_j \geq \pi_j$, $\forall j$.

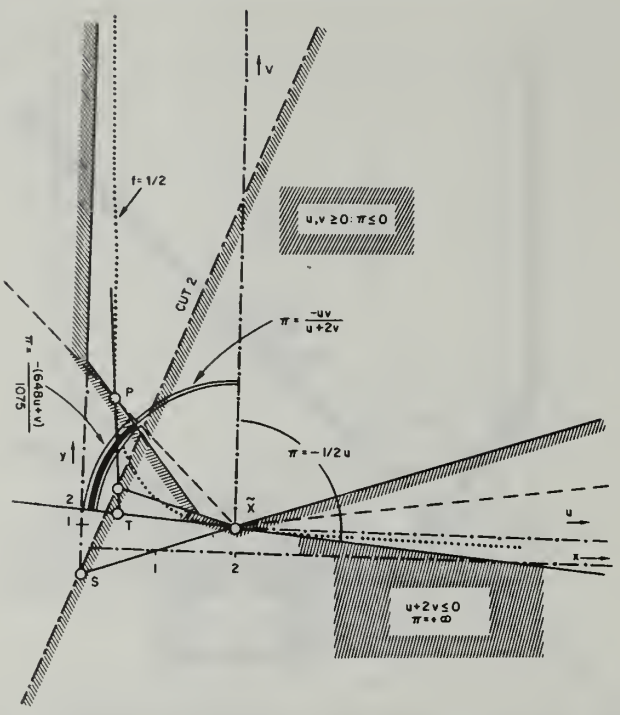


FIGURE 2d.—Here the set $\mathrm{lev}_{\pi_0}\pi$ is delimited by the curve.

$$f=\frac{1}{2}\ (\text{"above" } P)$$

the tangent line segment $P\ QT$ and the half-line through $T\tilde{x}$. One sees that $\pi=-v$ has disappeared and one now has

$$\pi=-\frac{1}{2}u,\ \pi=\frac{-uv}{u+2v},\ \pi=\frac{-(648u+v)}{1075}$$

and $\pi=+\infty$ in the four sectors defined in the illustration.

PROOF: Since $t_j\geq 0$, one has $\alpha_j t_j\geq\pi_j t_j$;
hence

$$\sum_j \alpha_j t_j\geq\sum_j \pi_j t_j\geq\pi_o \qquad \text{Q.E.D.}$$

COROLLARY 3.1.: The inequality

$$\sum_j \pi_j t_j\geq\pi_o$$

dominates Tuy's cut uniformly.

PROOF: Immediate from (3.10)

The "best" inequality can now be defined as one which dominates all others, i.e.

$$\bar{\pi}(u)=\inf_{\pi}\pi(u),$$

where π satisfies (3.8).

$\bar{\pi}$ is a convex gauge because it is defined by pointwise inf in a class of convex gauges.

In the above example $\bar{\pi}$ is easily seen to correspond to the choice

$$\bar{\pi}(u,v)=\max\left\{-\frac{1}{2}u,\ -v\right\},$$

for $u\geq 0$, $v\geq 0$.

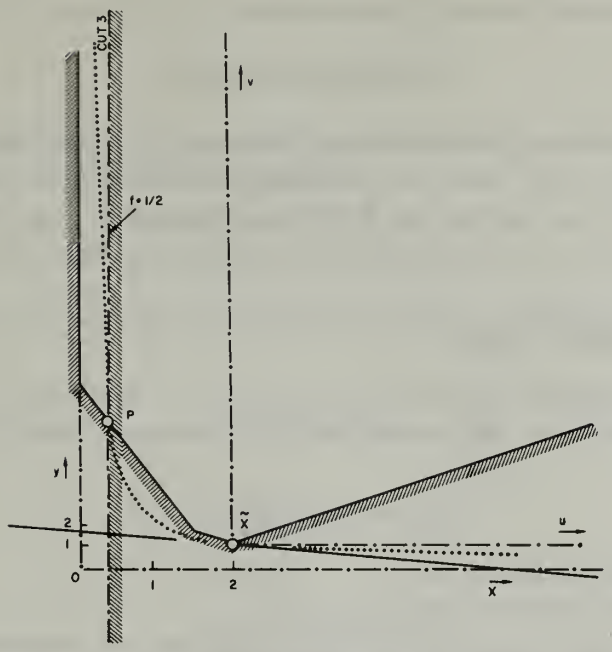


FIGURE 2e.—Illustration of the most stringent inequality: cut 3 is vertical through P.

We shall see in Section 4.1 that the gauge $\bar{\pi}$ is the *maximal convex extension* of π with respect to the set $S=R_I$. (See Remark $R1$ in Section 4.1 and cut 1 in Figure 2).

3.5. Subadditive Gauges Based on f and X

In his note [6] Tuy noticed that one could improve his cut by a conceptually simple argument. He characterizes a convex extension F of f with respect to X by:

$$\begin{aligned} &\text{(i)} \quad F(x) \leq f(x) && \forall x \\ &\text{(ii)} \quad F(x) = f(x) && \forall x \in X \\ &\text{(iii)} \quad F \text{ convex} \end{aligned} \tag{3.11}$$

This leads to the following "improvement" of the optimality condition (2.1):

$$X \subset \operatorname{lev}_{\bar{k}} f \Leftrightarrow X \subset \operatorname{lev}_{\bar{k}} F,$$

where the improvement stems from the property $\operatorname{lev}_{\bar{k}} f \subset \operatorname{lev}_{\bar{k}} F$.

Further results in this direction can be obtained within a subadditive framework:

(3.12a) a) π is a convex gauge with $\pi(0)=0$.

(3.12b) b) as (3.8b) where only those x which satisfy $x \epsilon X$ are considered.

REMARK: (3.12b) is thus a relaxation of (3.8b) and one has

$$(\operatorname{cl} \operatorname{lev}_{\pi_0} \pi) \cap X = [(\operatorname{lev}_k f) - \tilde{x}] \cap X; \tag{3.12c}$$

this π also generates a k-valid inequality; furthermore,

$$(\text{cl lev}_{\pi_o}\,\pi)\supset[(\text{lev}_k\, f)-\tilde{x}] \tag{3.13}$$

and from (3.8d), one sees that the inequality based on (3.12) will uniformly dominate the inequality based on (3.8). This result also dominates Tuy's extension (3.11): indeed (3.12) is based on the convexity of the *level sets* and therefore corresponds to a quasi-convex extension rather than the convex extension (3.11); moreover we do *not require f to be convex* (see also [2]).

3.6. The "Best" Subadditive Gauges

Finally one may bring an ultimate improvement in the definition of the gauge π which will yield (at least conceptually) k-valid inequalities which dominate uniformly all the others.
Define the set

$$S_\pi=\{x=x(t)\in X|\sum_j \pi_j t_j\leq\pi_o\}$$

and impose the following conditions on π:

(3.14a) a) π is a convex gauge, with $\pi(0)=0$

(3.14b) b) as (3.8b) where only those x which satisfy $x\epsilon S_\pi$ are considered.

REMARK: (3.14) is a relaxation of (3.12) because $S_\pi\subset X$.

LEMMA 5: The gauge π defined by (3.14) generates a k-valid inequality.

PROOF: Take $x\in X$ with $f(x)>k$. Suppose $x\in S_\pi$, then since

$$x=\tilde{x}+u=\tilde{x}+\sum_j e_j t_j,$$

one has

$$\pi_o<\pi(u)=\pi(\sum_j e_j t_j)\leq\sum_j \pi_j t_j$$

from (3.14b); but for all $x\in S_\pi$, one has by construction

$$\sum_j \pi_j t_j\leq\pi_o;$$

hence $x\notin S_\pi$ Q.E.D.

THEOREM 4: For every k-valid inequality

$$\sum_j \alpha_j t_j\geq\alpha_o(>0),$$

there exists a convex gauge π satisfying (3.14) and

$$\pi_j\leq\alpha_j,\ \forall j\in N$$

$$\pi_o=\alpha_o$$

PROOF: Define

$$S_\alpha=\{x\in X|\sum\alpha_j t_j\leq\alpha_o\};$$

by hypothesis S_α contains no point x with $f(x)>k$. Setting $\pi_o=\alpha_o$ and $\pi(e_j)=\alpha_j$ (with $\pi(0)=0$) one defines a gauge π, which is a hyperplane, and which satisfies (3.14); in general, there may exist a way to modify the above hyperplanar gauge into a convex gauge π, so that one may have the more general relation $\pi_j\leq\alpha_j$, $\forall j\in N$ Q.E.D.

REMARK: The above theorem indicates that every supporting hyperplane of the convex hull

$$\text{conv } \{x \in X | f(x) \geq k\}$$

may be generated by a subadditive gauge π. Thus, in principle, the subadditive gauge approach can produce the most stringent k-valid inequalities.

3.7. Example

We now illustrate the construction of k-valid inequalities as described in the Sections 3.4, 3.5, and 3.6.

Let

$$f(x, y) = (xy)^{-1}$$

and

$$X: 4x+y \geq 22/3, \qquad x, y \geq 0$$
$$x+y \geq 3$$
$$x-y \leq 1 \text{ (see Figure 2).}$$

3.7.1. "Tuy" and "Cut 1," Choose $\tilde{x}=(2, 1)$ and define the gauge π as in the example of Section 3.4. One obtains the following coefficients:

$$\pi_1 = \pi(-1, 1) = (\hat{\lambda})^{-1} = 1$$

$$\pi_2 = \pi(1, 1) = \max\left\{-\frac{1}{2}, -1\right\} = -\frac{1}{2}$$

For the $\left(k=\frac{1}{2}\right)$—valid inequalities, one has $(\pi_o=1)$:

$$\text{Tuy: } t_1 \geq 1, \text{ i.e. } -\frac{1}{2}x+y \geq 1$$

$$\text{Gauge } \pi\text{: } t_1 - \frac{1}{2}t_2 \geq 1 \text{ i.e. } x-y \geq 1$$

π is the maximal convex extension with respect to S_1 (see end of Section 3.4 and Section 4.2).

3.7.2. "Cut 2." The above cuts are based solely on $\text{lev}_{1/2} f$; the following gauge now takes the set X into consideration (3.12):

I) in the halfspace $\frac{1}{2}u+v \geq 0$:

IA) $$u \leq 0,\ 3u+v \geq 0\text{: } \pi(u, v) = \frac{-uv}{(u+2v)}$$

$$\pi(0, 0) = 0$$

IB) $$(u \leq 0 \text{ and } 3u+v \leq 0) \text{ or } u \geq 0\text{: } \pi(u, v) = \frac{-1}{1075}(648u+v)$$

II) $$\text{for } \frac{1}{2}u+v<0: \pi(u, v)=+\infty$$

Thus one has

$$\pi_1=\frac{647}{1075}$$

$$\pi_2=-\frac{1}{2}$$

3.7.3. "Cut 3." Finally we also show (cut 3) the inequality corresponding to the convex hull described in the final remark of section 3.6.

4. GAUGE EXTENSIONS

In this section, we apply the general methodology developed in [2] to characterize some f and X related gauge functions (see Section 3.5). These gauges are described as *(maximal) convex extension with respect to the set* X of the f related gauge functions of Section 3.4.

It is also shown that the latter are (maximal) convex extensions of the Tuy-type gauge functions given in Corollary 2.1. Thus the concept of (maximal) convex extensions of a convex function with respect to a given set appears here as a fundamental tool.

We then derive a result indicating that the *maximal convex gauge extension* corresponds in fact to a convex gauge based on the *maximal quasi-convex extension of* f; it therefore dominates all previously known results of this type, in particular those due to Tuy [6].

Finally we introduce the notion of *polaroid gauge functions* (see also [3]). It is shown that, when the level set gauge is differentiable ($\nexists u \neq 0$), the polaroid approach offers an analytical method to determine the *maximal convex gauge extension.*

4.1. Convex Gauge Extensions

Consider the convex gauge π, defined on R^n; and let S be a subset of R^n, with non-empty interior (Int $S\neq\emptyset$).

Since π is a convex gauge, the set epi π is convex and it can be represented by the intersection of all halfspaces H which support epi π: $H(y)=\{(h, u)|h\geq H_y u,\ u\in R^n\}$ with $h=H_y y=\pi(y)$ and $\pi(u)$ can now be represented by

(4.1) $$\pi(u)=\sup_{y\in R^n}\{H_y u\}.$$

The *maximal convex gauge extension* Π *of* π *with respect to* S is constructed by relaxing (4.1) where only certain points y in the sup are selected:

DEFINITION 3:

(4.2) $$\Pi(u)=\sup_{y\in \text{Int } S}\{H_y u\}$$

PROPERTIES:

P1) One has $\pi(u)\geq\Pi(u)$, $\nexists u$ which implies dominance of the inequalities generated from π by those stemming from Π.

P2) Π is convex

(4.3)

P3) Π is a gauge

(4.4)

Proof: $$\Pi(\lambda u)=\sup_{y\in \operatorname{Int} S}\{H_y\lambda u\}=\lambda \sup_{y\in \operatorname{Int} S}\{H_y u\}=\lambda\Pi(u), \text{ for } \lambda\geq 0$$ Q.E.D.

P4) $\operatorname{lev}_{\Pi_o}\Pi=\operatorname{lev}_{\pi_o} MQC\pi\supset\operatorname{lev}_{\pi_o}\pi$, for $\Pi_o=\pi_o\geq 0$ where $MQC\pi$ denotes the maximal quasi-convex extension of π with respect to the set S (see [2] for further details on MQC functions).

PROOF: We need show that the set $\operatorname{lev}_{\pi_o}\Pi$ is the maximal convex extension (with respect to S) of the set $\operatorname{lev}_{\pi_o}\pi$. For each $y\in$ Int S, one has by construction: $\Pi(y)=\pi(y)$ and therefore also $\Pi(\lambda y)=\pi(\lambda y)$, $\forall\lambda\geq 0$. Now if π_o is such that $\operatorname{lev}_{\pi_o}\pi\cap$ Int $S=\emptyset$, we simply have $\Pi\equiv 0$, by definition. Otherwise consider the collection of halfspaces

$$G(y)=\{u\in R^m \mid H_y u\geq\pi(y)=(H_y y)\}.$$

One has

$$\operatorname{lev}_{\pi_o}\Pi=\operatorname{cl}\bigcap_{y\in(\operatorname{Int} S\cap \operatorname{lev}_{\pi_o}\pi}G(y),$$

for $\pi_o\geq 0$ which is (by definition) the maximal convex extension of the set $\operatorname{lev}_{\pi_o}\pi$ with respect to S (see [2], Theorem 2). Q.E.D.

P5) By construction of the level set gauge π (3.8), one has $\operatorname{lev}_{\pi_o}\pi=[(\operatorname{lev}_k f)-\tilde{x}]$, where $k\geq f(\tilde{x})$, $\tilde{x}\in S$; thus one also has $\operatorname{lev}_{\pi_o} MQC\pi=[(\operatorname{lev}_k MQCf)-\tilde{x}]$; since both MQC functions are obtained from the maximal convex extension of the same sets;

Thus from the above property P4, we observe that the convex gauge extension Π is the gauge belonging to $MQCf$, not f; it therefore brings about "better" optimality conditions for the program (2.1).

REMARKS: There are two instances for which an extension process is called for in our approach:

R1) for the definition of the convex gauge π along rays where the function f is unbounded ($S=R_I$) as explained at the end of Section 3.4.

R2) for the construction of a gauge which takes the feasible region X into account $S=X-\tilde{x}$

Thus one may naturally choose the following set S in Definition 3:

(4.5) $$S=[\{x\in -X \mid (x)\geq k\}f\tilde{x}]$$

in order to combine both aspects R1 and R2.

4.2. Polaroid Gauges

We now use a generalized concept of polarization introduced in [1] to construct subadditive gauges. Consider a function $\Phi(u;p)$: $R^m\times P\to R$ and a set P; assume $\Phi(\cdot;p)$ to be a convex gauge $\forall p\in P$ and define the *polaroid gauge* (see [1, 3]):

(4.6) $$\Pi(u)=\sup_{p\in P}\Phi(u;p)$$

LEMMA 6: $\Pi(u)$ is a convex gauge

PROOF (omitted): See Lemma 7 and 8 of [3]

COROLLARY 6.1.: The inequality

$$\sum_j \Pi_j t_j \geq \Pi_o$$

is k-valid with

$$\Pi_o = \min_{u \in S} \sup_{p \in P} \Phi(u; p) \tag{4.7}$$

EXAMPLE: Let $\hat{\pi}(u)$ be a level set gauge defined by (3.10) and assume $\hat{\pi}$ differentiable for all $u \neq 0$ with $\hat{\pi}(u) < +\infty$.
Define

$$\Phi(u; p) = \hat{\pi}(u) + (p-u)\nabla\hat{\pi}(u) \tag{4.8}$$

Then choose $P = S \subset R^n$ to obtain for each $u \in S$:

$$\Pi(u) = \sup_{p \in P} \Phi(u; p) = \Phi(u; u) = \hat{\pi}(u)$$

Thus

$$\Pi_o = \min_{u \in S} \hat{\pi}(u) = \hat{\pi}_o,$$

the same value as in Corollary 2.1.

Note that $\Pi(u)$ is in general different from $\pi(u)$ for $u \notin S$. Furthermore, one can show (see [2], Lemma 12) that the gauge Π constructed here is, in fact, the *maximal convex extension of* $\hat{\pi}$ *with respect to* S.

Note that this extension is obtained from a polarization of the level set gauge $\hat{\pi}$, not of the objective function f. Although one has, by definition, $\operatorname{lev}^{\wedge}_{\hat{\pi}_o} \hat{\pi} = [(\operatorname{lev}_k f) - \tilde{x},]$ the polaroid gauges derive from $\hat{\pi}$ and f are quite different in general.

REMARKS: 1) If $\hat{\pi}(u)$ is not differentiable $\forall u \neq 0$, the gradient $\nabla\hat{\pi}(u)$ can be replaced by the subdifferential $\partial\hat{\pi}(u)$ in (4.8) since π is convex: i.e.

$$\Phi(u; p) = \hat{\pi}(u) + \sup_{\nabla\hat{\pi} \in \partial\hat{\pi}(u)} (p-u)\nabla\pi \tag{4.9}$$

2) In principle, the determination of the coefficients Π_j, and Π_o are obtained from an *auxiliary program* i.e.

$$\Pi_j = \Pi(e_j) = \sup_{p \in P} \Phi(e_j; p) \tag{4.10a}$$

$$\Pi_o = \min_{u \in S} \sup_{p \in P} \Phi\left(\sum_j e_j t_j; p\right) \tag{4.10b}$$

with

$$u = \sum_j e_j t_j$$

$$t_j \geq 0, \ \forall j.$$

The nature and degree of difficulty of these programs (4.10) largely depend upon the set P and the form of the function Φ

Since there are n coefficients Π_j one will try to keep (4.10a) as simple as possible.

EXAMPLES: In the following two examples, the program (4.10a) reduces to the explicit determination of the largest number in a finite list:

1) P finite $=\{p_1, p_2, \ldots, p_P\}$

i.e. $$\Phi(u; p)=\Phi_p(u),\ p\in P$$

2) P may also be a list of finite index sets

$$\mathrm{P}=\{P_1, P_2, \ldots P_T\}$$

with the following convex gauges $\forall \tau\in T$:

$$\Phi_\tau(u; p),\ p\in P_\tau$$

For instance $$\Phi_\tau(u)=\sum_{p\in P_\tau}\Phi_p(u),$$

and the corresponding polaroid gauge becomes:

$$\Pi(u)=\max_{\tau\in T}\sum_{p\in P_\tau}\Phi_p(u).$$

We may now analyse Example 3.7 and Figure 2. Starting from Tuy's cut (and the corresponding level set gauge which has $\hat{\pi}(u, v)=0, \forall u, v\geq 0$ and

$$\hat{\pi}(u, v)=\frac{-uv}{(u+2v)}\ \text{otherwise}\Big)$$

the cut 1 is obtained by polarization (4.8) with $S_1=\{(u, v)|u\leq 0, \text{ or } v\leq 0\}$; this corresponds to remark $R1$, the end comments of section 4.1, and to the end comments of section 3.4.

For the corresponding polaroid gauge one has: $\Pi(u)=+\infty$ in the halfspace $u+2v<0$; otherwise the halfspace $u+2v\geq 0$:

$$\Pi(u, v)=\sup_{\substack{\mu\leq 0\\ \text{or}\\ v\leq 0}}(\mu+2v)[-(2v^2u+\mu^2v)]=\begin{cases}\dfrac{-uv}{(u+2v)}, & \text{for } u\leq 0 \text{ or } v\leq 0\\ \max\left\{-\dfrac{1}{2}u, -v\right\} & \text{for } u, v\geq 0\end{cases}$$

Similarly for the cut 2, with

$$S_2=\left[\left\{x\in X|f(x)\geq\frac{1}{2}\right\}-\tilde{x}\right].$$

BIBLIOGRAPHY

[1] Burdet, C. A., "Polaroids: A New Tool in Non-Convex and in Integer Programming," Naval Research Logistic Quarterly, 20, 13–24 (1973).

[2] Burdet, C. A., "Convex and Polaroid Extensions," WP 73–21, Faculty of Management Sciences, University of Ottawa (1973).

[3] Burdet, C. A., "On the Algebra and Geometry of Cuts," WP 74–8, Faculty of Management Sciences, University of Ottawa (1974).

[4] Burdet, C. A., "On Polaroid Intersections," Mathematical Programming in Theory and Practice, P. Hammer and G. Zoutendijk, eds., pp. 365–387 (North Holland, 1974).

[5] Burdet, C. A., "The Facial Decomposition Method," Operations Research Quarterly, *24*, 459–463 (1973).

[6] Tuy, Hoang, "Concave Programming Under Linear Constraints," (Russian) *Doklady Akademii Nauk SSSR*, 1964. English translation in Soviet Mathematics, 1437–1440 (1964).

[7] Johnson, E. L., "A Group Problem for Mixed Integer Programs," to appear in Mathematical Programming (1975).

[8] Rockafellar, R. T., "Convex Analysis," (Princeton University Press, 1970).

CONVEX AND POLAROID EXTENSIONS

Claude-Alain Burdet

SYSTEMATHICA Consulting Group Ltd.
Pittsburgh, Pennsylvania

ABSTRACT

In an effort towards a comprehensive and unified theory, this note presents some new results in the area of non-convex programming within the framework of convex (sets and function) analysis.

The entire study is primarily devoted to the development of useful tools for extreme point programs (such as concave or integer programs).

0. BACKGROUND

This paper presents several new ideas in convex analysis to come to grips with non-convex programming problems. It lies in the mainstream of developments in the theory of polar sets [7], generalized polars [1] and polaroids [2]; although some results heavily rely on concepts introduced in [2 and 3] the paper is self-contained. The line of research pursued here is related to the following general *constrained optimization* problem:

Consider the non-linear program

$$\text{maximize } f(x), \text{ subject to } x \in X \subset \mathbb{R}^n$$

One has immediately the following paraphrase of the global optimality property which may be interpreted as a *necessary and sufficient optimality condition*

$$\left.\begin{array}{l}\bar{x} \in X, \text{ with } \bar{k} = f(\bar{x}) \\ \text{is globally optimal}\end{array}\right\} \Leftrightarrow \text{lev}_{\bar{k}} f \supset X.$$

The object of this note is to establish weaker *sufficiency conditions* for the case where both $\text{lev}_{\bar{k}} f$ and X are *convex*.

In his short note [5], Tuy proposed some interesting ideas which can be applied to this kind of problem, particularly when X is *polyhedral*. This paper is an attempt to produce more powerful results (both theoretically and practically) and to provide some elements of mathematical structure in this area.

In a first section we introduce the concept of *convex extension of a set with respect to another*. It includes a pointwise definition, a study of the question of *maximality*, and an equivalent definition by means of hyperplanes.

This concept is applied to define *convex* and *quasi-convex extensions of a given convex function f* with respect to a (convex) set (Sections 2 and 3); we also construct a *maximal convex extension of f*, which is shown to have the same properties as the extension described geometrically by Tuy [5]; furthermore, a dominance theorem is established.

Section 4 briefly reviews an application of these extensions to the non-convex programming problem area.

The next section features another type of extension called *polaroid extension* which possesses the advantage (over the previous two types of extension) of being more computationally tractable.

Section 6 handles the special case where f is assumed differentiable; an analytical definition of the maximal convex extension of f is presented; it is also imbedded in a family of polaroid extensions. The end of this section is devoted to a brief description of the *quadratic* case [1], where the general results derived in the previous sections assume a special form; parallel studies [1, 4 and 8] develop the corresponding algorithmic implementations.

We then examine in greater depth the application of the present results to non-convex programs; it is shown how a valid cut can be constructed from an extension.

In the second part of Section 7 a new type of cut based on a *cutting polaroid* is established: it combines cutting planes and polaroids to produce a "deep" cut where the two theories interact.

Finally a last section indicates some further improvements which can be obtained when the feasible set is polyhedral.

To conclude this introductory tour it should be pointed out that little effort is made here to investigate *every* aspect of possible interest in the theory of polaroids. We have tried instead to adopt a line of development which seems most promising as far as implementation of new ideas is concerned. This has been done with the double intent of obtaining early results of tangible interest and of attracting others to the development of an open area.

1. SET DEFINITION AND PROPERTIES

Consider a closed convex set C in $\mathbb{R}^n$ and a convex set D with non-empty interior (i.e., $\mathbb{R}^n = \text{aff } D$).

DEFINITION 1: A *convex extension E of C with respect to D* is defined to satisfy:

a) E is convex

b) $E \supset C$

c) $(E \cap \text{Int } D) \subset (C \cap \text{Int } D)$

Note: b) and c) imply $(E \cap \text{Int } D) = (C \cap \text{Int } D)$.

LEMMA 1: E is *closed*.

PROOF: We show that the complement of E is open; take $x_1 \notin E$, then there exists $x_2 \in E$ such that

$$x_3 = \hat{\lambda} x_1 + (1-\hat{\lambda}) x_2, \ \hat{\lambda} \in (0, 1]$$

satisfies

$$x_3 \in \text{Int } D \text{ but } x_3 \notin C.$$

C is closed, hence there exists an open ball $B(x_3)$ such that

$$B(x_3) \subset D \text{ but } B(x_3) \cap C = \emptyset.$$

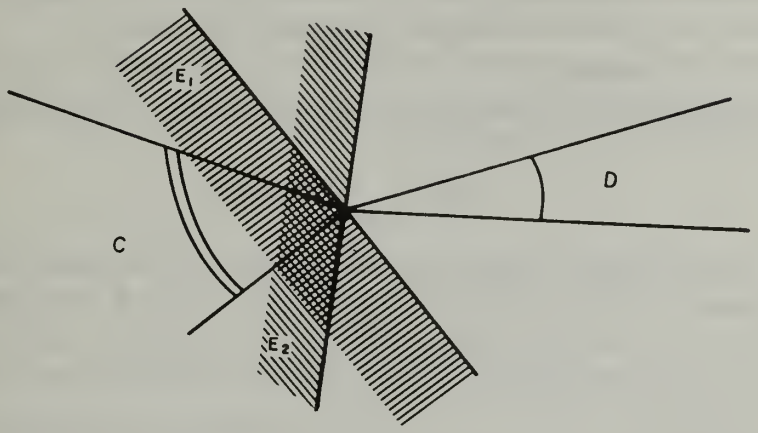
Since $\hat{\lambda} \neq 0$, we may define the following open neighborhood of x_1

$$U(x_1) = \{x = \hat{\lambda}^{-1}[y - (1-\hat{\lambda})x_2] \mid y \in B(x_3)\};$$

and $\forall x \in U(x_1)$ there exists $y \in$ Int D with $y \notin C$ hence $x \notin E$, i.e., $U(x_1) \cap E = \emptyset$. Q.E.D.

The following example shows that E need not be unique

EXAMPLE 1:



DEFINITION 2: A convex extension E is said *maximal* if $E \supset E'$ for any convex extension E' of C with respect to D).

LEMMA 2: A convex extension E is *maximal* iff it contains all points x_1 such that $\forall x_2 \in C$, one has $x_3 \in$ Int $D \Rightarrow x_3 \in C$, with $x_3 = \lambda x_1 + (1-\lambda)x_2$, $\lambda \in [0, 1]$.

PROOF: From definition 1, one sees that every point x_1 of a convex extension satisfies the above; if a convex extension E contains all such points, it contains every extension and is therefore, maximal.

Conversely, we show that there always exists a convex extension which contains x_1; it then follows from the maximality of E that $x_1 \in E$. Construct the set

$$V(x_1) = \{x \mid x = \lambda x + (1-\lambda)y,\ \lambda \in [0, 1],\ y \in C\}.$$

Clearly V is convex; from the hypothesis one also has

$$V(x_1) \cap \text{Int } D = C \cap \text{Int } D. \qquad \text{Q.E.D.}$$

Example 1 shows that the set of *all* points x_1 (Lemma 2) need not be convex (and therefore, not an extension E according to definition 1).

THEOREM 1 (existence and uniqueness): *If* $(C \cap \text{Int } D) \neq \emptyset$, there exists exactly one maximal extension E.

PROOF (existence): The set Q of points x_1 satisfying the conditions of Lemma 2 satisfies b) and c) of Definition 1; it remains to show convexity.

Let $x_1^3 = \lambda x_1^1 + (1-\lambda)x_1^2$, $\lambda \in [0, 1]$ where $x_1^1, x_1^2 \in Q$.

First we note that if the (convex) segment $x_1^1 x_1^2 (= S)$ intersects C then $x_1^3 \in Q$ because x_1^1 and $x_1^2 \in Q$ (see Lemma 2); furthermore: $\forall y \in$ Int D, $y \in C$ iff $y \in E$. Thus, we may restrict our-

selves to the case where the segment S does not intersect C. Construct the set $V(x_1^1)$ and $V(x_1^2)$ (see Lemma 2); define $\hat{y} \in bd\ C$ with $\hat{y} = \hat{\mu} x_1^3 + (1-\hat{\mu}) y$, $\hat{\mu} \in (0, 1)$ for any $y \in C \cap \text{Int}\ D$; it is sufficient to show $\hat{y} \notin \text{Int}\ D$, implying $x_1^3 \in Q$.

One has $\hat{y} \in \text{cl}\ [V(x_1^1) - C] \cap \text{cl}\ [V(x_1^2) - C]$, because $S \cap C = \emptyset$; but $[V(x_1^i) - C] \cap \text{Int}\ D = \emptyset$, hence also $\text{cl}\ [V(x_1^i) - C] \cap \text{Int}\ D = \emptyset$ and therefore, $\hat{y} \notin \text{Int}\ D$.

(Uniqueness): For any two maximal extensions E, E' one must have $E \subset E'$ and $E' \subset E$ hence $E = E'$.

Q.E.D.

Theorem 1 provides a pointwise definition of the maximal convex extension of C with respect to D. We now turn to an equivalent definition, by means of halfspaces.

THEOREM 2: The maximal convex extension E of C with respect to D is the closed intersection $\forall y \in (C \cap \text{Int}\ D) \neq \emptyset$ of all halfspaces $H(y) = \{x | h_y x \leq h_0\}$ corresponding to hyperplanes $(h_y x = h_0$ which support C at y (i.e., $h_0 = h_y y$ and $h_y x \leq h_0$, $\forall x \in C$):

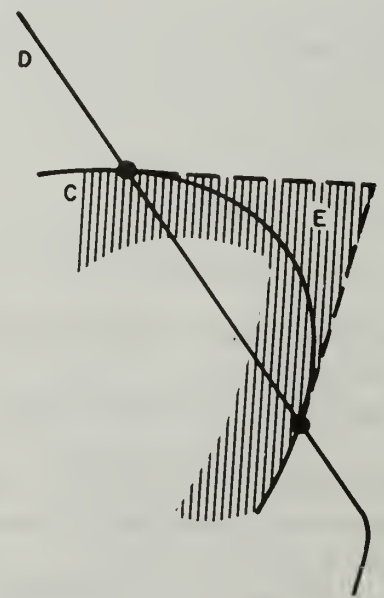
$$E = \text{Cl} \bigcap_{y \in C \cap \text{Int}\ D} H(y)$$

PROOF: E clearly contains C and is convex; furthermore, $(E \cap \text{Int}\ D) = (C \cap \text{Int}\ D)$. Hence E is a convex extension. We next show that E contains every point x_1 (see Lemma 2); take any x_1, i.e., $\forall x_2 \in C$ one has $x_3 \in \text{Int}\ D \Rightarrow x_3 \in C$. Suppose there exists a supporting hyperplane at $x'_2 \in C \cap \text{Int}\ D$ separating x_1 from C; then there exists no $x'_3 = \lambda x_1 + (1-\lambda) x'_2$, $\lambda \in [0, 1]$ with $x'_3 \in C$; however, there is an open ball $B(x'_2) \subset D$ so that $x'_3 \in B(x'_2)$; this contradicts the property of x_1 stipulated in Lemma 2. Thus, there is no such separating support and $x_1 \in E$. Maximality of the set of points x_1 (Lemma 2) now implies the reverse inclusion; hence both definitions characterize the same set E.

Q.E.D.

Note that we have to define E as the closure of this intersection of halfspaces because, in general, this intersection may not be closed as shown in Example 2.

EXAMPLE 2



It may also be useful to remark that this concept of extension of a set with respect to another can be generalized to non-convex sets.

2. CONVEX EXTENSIONS

Consider now the real-valued function f; let X be a closed convex subset of $\mathbb{R}^n$ and assume $X \subset \operatorname{dom} f$ and $\operatorname{Int} X \neq \emptyset$.

LEMMA 3: Consider a convex set $\Delta \subset (\mathbb{R}^n \times \mathbb{R})$ and define

$$f(x) = \inf_{(x,\, r) \in \Delta} (r, +\infty)$$

then f is *convex.*

PROOF: Take $x = \lambda x_1 + (1-\lambda) x_2$. We show $f(x) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$, $\lambda \in [0, 1]$. By convexity of Δ one has:

$$f(x) = \inf_{(x,\, r) \in \Delta} r \leq \lambda r_1 + (1-\lambda) r_2 \text{ for any } r_1, r_2$$

such that (x_1, r_1) and $(x_2, r_2) \in \Delta$; hence, also for $f(x_1) = \inf r_1$ and $f(x_2) = \inf r_2$. Q.E.D.

The converse of Lemma 3 is, of course, also true i.e., the *epigraph*

$$\Delta = \operatorname{epi} f = \{(x, r) \mid x \in X,\ r \in \mathbb{R},\ r \geq f(x)\}$$

is *convex.*

Define the (convex) cylinder D over the convex set X

$$D = \{(x, \delta) \mid x \in X,\ \delta \in \mathbb{R}\} \subset \mathbb{R}^{n+1}$$

DEFINITION 3: A convex extension CF of the function f with respect to the (closed convex) set X is a function whose epigraph G is the convex extension of the set epi f with respect to D i.e.,

$$CF(x) = \min \{+\infty, \varphi \mid (x, \varphi) \in G\}.$$

In mathematical programming, we are really only interested in the values of f over the feasible set X; the next Lemma shows that f and CF are "essentially" identical on X, i.e., that one may replace f by its extension CF.

LEMMA 4: $\forall x \in \operatorname{Int} X$, one has $CF(x) = f(x)$.

PROOF: For $x \in X$ one has $(x, \varphi) \in D$, $\forall \varphi \in \mathbb{R}$; furthermore, from Definition 1

$$(\operatorname{Int} D \cap \operatorname{epi} f) = (\operatorname{Int} D \cap \operatorname{epi} CF)$$

holds true. Hence, $\forall x \in \operatorname{Int} X$,

$$CF(x) = \min \{+\infty, \varphi \mid (x, \varphi) \in \operatorname{epi} CF\} = \min \{+\infty, \varphi \mid (x, \varphi) \in \operatorname{epi} f\} = f(x).$$ Q.E.D.

We now examine the discrepancies between f and CF, i.e., the special cases where the two are not rigorously identical.

LEMMA 5: The *convex extension* of a convex function on X is a *continuous convex* function.

PROOF: Convexity follows from convexity of epi CF and Lemma 3. Furthermore, the set epi CF is closed (Lemma 1) implying continuity of the min function. Q.E.D.

Note that Lemma 5 does not require continuity of f; and, if f is discontinuous, one may have $f(x) \neq CF(x)$ for some $x \in \operatorname{bd} X$.

COROLLARY L5.1: If f is continuous on X then $f(x)=CF(x)$, $\forall x\in X$.

PROOF: Take $x\in \text{bd } X$; continuity of CF and f, with $f(x)=CF(x)$, $\forall x\in \text{Int } X$ implies equality on the boundary also. Q.E.D.

Maximality was introduced in Section 1 because, as we shall see below, it introduces in a natural way the most "desirable" extensions, i.e., which *dominate* the others and yield better sufficient conditions. If one considers the *maximal* convex extension in the definition 3, then the resulting function MCF is called *maximal* convex extension of f; and one has:

THEOREM 3: For any real-valued convex function f, f' with $f'(x)=f(x)$, $\forall x\in X$ one has

$$f'(x)\geq MCF(x),\ \forall x\in \mathbb{R}^n.$$

PROOF: Maximality of epi MCF with respect to D implies epi $f'\subset$ epi MCF. Q.E.D.

Note that $\forall x\in \mathbb{R}^n$ one has $MCF(x)\leq f(x)$ so that MCF can only become $+\infty$ when and where f has the same property.

In the special case where f is not *continuous* on X (discontinuities may only occur on bd X), the epigraph should be replaced by its *closure*: cl epi f.

The following immediate corollary relates the present construction to the concept envisaged by Tuy in [5]:

COROLLARY T3.1: A convex function $g(x)$ satisfies (i, ii, iii) below iff $g(x)=MCF(x)$.

(i) $g(x)=f(x)$, $\forall x\in X$

(ii) $g(x)\leq f(x)$

(iii) (maximality): $g(x)\leq f'(x)$ for any convex function satisfying (i), (ii).

PROOF: From Theorem 3 we know that MCF satisfies the above conditions (i)–(iii); in particular $MCF(x)\leq g(x)$ furthermore, MCF is convex, hence from (iii) one must have $g(x)\leq MCF(X)$. Q.E.D.

Thus, one has: The maximal convex entension of the function f with respect to X is *unique*.

3. QUASI-CONVEX EXTENSIONS

In many ways, and particularly in the context of the optimality condition described in Section 0, convex extensions appear to be too restrictive; indeed only quasi-convexity is really required, and we now turn to the definition of the corresponding generalizations.

DEFINITION 4: f is a quasi-convex function if and only if its *level sets* $\text{lev}_\varphi f=\{x|f(x)\leq\varphi\}$ are *convex*.

DEFINITION 5: A *quasi-convex extension* QCF of the function f with respect to the (convex) set X is a function whose level set L_φ is a convex extension of $\text{lev}_\varphi f$ with respect to X; by definition one has

$$QCF(x)=\inf\{\varphi, +\infty|x\in L_\varphi\}.$$

It may happen here that $\text{lev}_\varphi f$ is not closed (when f is discontinuous); one then uses cl $\text{lev}_\varphi f$ to define the extension.

LEMMA 6: $\forall x\in \text{Int } X$, one has $QCF(x)=f(x)$.

PROOF: By Definition 1, one has

$$\text{Int } X\cap \text{lev}_\varphi f=\text{Int } X\cap L_\varphi,\ \forall\varphi\in f(X)$$

thus, for $x\in \text{Int } X$ one has $f(x)=\varphi\Leftrightarrow QCF(x)=\varphi$. Q.E.D.

LEMMA 7: The quasi-convex extension of a quasi-convex function on X is a *continuous* quasi-convex function.

PROOF: As for Lemma 5, this follows from convexity and closedness of the convex extension L_φ. Q.E.D

Here again f is not required continuous and one may have $f(x) \neq QCF(x)$, for $x \in$ bd X (see remark after Lemma 5). If L is the *maximal* extension $\forall\varphi$ one obtains the *maximal* quasi-convex extension *MQCF* with the following property:

THEOREM 4: For any real-valued quasi-convex functions f, f' with $f'(x) = f(x)$, $\forall x \in X$, one has $f'(x) \geq MQCF(x)$, $\forall x \in \mathbb{R}^n$.

PROOF: By construction $\text{lev}_\varphi\, MQCF$ is maximal $\forall\varphi$ hence,

$$\text{lev}_\varphi f' \subset \text{lev}_\varphi\, MQCF \text{ must hold } \forall\varphi.$$

Q.E.D.

Clearly, since a convex extension is also quasi-convex, one may consider both extensions of a convex function; one then has the following *uniform dominance* theorem.

COROLLARY T4.1: $\text{lev}_\varphi f \subset \text{lev}_\varphi\, MCF \subset \text{lev}_\varphi\, MQCF$, $\forall\varphi$ $f(x) \geq MCF(x) \geq MQCF(x)$, and equality holds throughout on Int X. An application of this main result is presented in the next section.

In conclusion of this first part of the paper, we observe that quasi-convex extensions provide the proper theoretical foundations (rather than convex extensions) for the investigation of optimality conditions within the framework of convex analysis. It should also be pointed out that, in spite of its trivial derivation, the above *dominance* theorem is of great practical importance; the reader may also convince himself that the quasi-convex extension *MQCF* has a graph different from *MCF*.

4. (QUASI-) CONCAVE PROGRAMMING

Let us consider our initial non-linear program (see Section 0): maximize $f(x)$ subject to $x \in X \subset \mathbb{R}^n$, where X is assumed (closed) convex and f (quasi-) convex. Thus, the *inclusion test* $\text{lev}_\varphi f \supset X$ required by the sufficiency conditions deals with two convex sets and we can use the Uniform Dominance Theorem T4.1 to obtain *better* optimality conditions than the above. We use the word "better" to indicate that, for whatever method used to actually perform an inclusion test, it is reasonably clear that such a test is more conveniently performed on $Q_1 \supset X$ rather than on $Q_2 \supset X$ when $Q_1 \supset Q_2$. The following theorem shows that sufficiency is indeed maintained.

THEOREM 5: Assume continuity of f on X; then one has

$$X \subset \text{lev}_k f \Leftrightarrow X \subset \text{lev}_k\, MCF \Leftrightarrow X \subset \text{lev}_k\, MQCF$$

and

$$f(x) = MCF(x) = MQCF(x),\ \forall x \in X.$$

PROOF: Immediate from Lemma 5 and 7.

REMARK 1: The possibility of improving the sufficiency inclusion test by considering a convex *extension* has first been pointed out by Hoang Tuy in [5]; the equivalence of his description of a convex extension with ours is shown in Theorem 3. The introduction of a quasi-convex extension serves the double purpose of providing yet a better condition when f is convex and also a possibility to handle the quasi-convex case in a similar manner.

REMARK 2: Since the condition $\text{lev}_{\bar{k}} f \supset X$ is necessary and sufficient for optimality of $\bar{x} \in X$ (with $f(\bar{x}) = \bar{k}$), Theorem 5 describes a variety of necessary and sufficient optimality conditions; it should be noted however that while extensions have the property of improving (weakening) the sufficiency conditions, this has an adverse effect on the necessary aspect since one wants the strongest possible necessary condition. On the other hand, our assumptions on X and f renders this "necessary" side of the question uninteresting because it has been thoroughly developed in the literature.

5. POLAROID EXTENSIONS

The definitions of (quasi-) convex extensions do not lend themselves easily to a straightforward algorithmic implementation because they are geometric rather than analytical. One may also encounter practical difficulties when f is not continuous on X. We now introduce the concept of *polaroid extensions* to remedy this situation.

DEFINITION 6 [2]: Given $k \in \mathbb{R}$, the *polaroid set* $X^*(k)$ corresponding to the set X *with respect to the function* φ: $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is defined by

$$X^*(k) = \{y \mid \varphi(x; y) \leq k, \ \forall x \in X\} \subset \mathbb{R}^n.$$

DEFINITION 7: The function φ is said to *polarize* f on X if $\forall x \in X$ one has $f(x) = \varphi(x; x)$. The polarization is said *proper* when φ satisfies $\varphi(x; y) \leq \max \{f(x), f(y)\}$.

THEOREM 6: If $\forall x \in X$, the function $\Phi(y) = \varphi(x; y)$ is *quasi-convex* (in y) then the polaroid $X^*(k)$ is *convex* $\forall k$.

PROOF: Omitted (see [2, 3]).

In the sequel, we shall always assume that φ satisfies the hypothesis of Theorem 6, as we only make use of *convex polaroids*.

We can now apply the above concepts to our present objective of constructing a (quasi-) convex extension of f; to this effect we now consider the set $(X \cap \text{lev}_k f)$ and its polaroid $P(k)$:

LEMMA 8: Under proper polarization, the polaroid $P(k) = (X \cap \text{lev}_k f)^*(k)$ is a *convex extension* of the set $\text{lev}_k f$ with respect to X.

PROOF: a) convexity follows from Theorem 6

b) $P(k) \supset \text{lev}_k f$: one has $P(k) = \{y \mid \varphi(x; y) \leq k, \ \forall x \in (X \cap \text{lev}_k f)\}$; take $y \in \text{lev}_k f$, $\forall x \in (X \cap \text{lev}_k f)$ then $\max \{f(x), f(y)\} \leq k$ holds true. Hence, since the polarization φ is proper,

$$\varphi(x; y) \vee \max \{f(x), f(y)\} \leq k, \text{ i.e., } y \in P(k).$$

c) Take $z \in X \cap \text{lev}_k f$, with $f(z) = k$; the set $Q = \{y \mid \varphi(z; y) \leq k\}$ is convex since $\varphi(z, y)$ is assumed quasi-convex in y; furthermore, the following insertions hold $Q \supset P(k) \supset \text{lev}_k f \supset (X \cap \text{lev}_k f)$; moreover, $f(z) = \varphi(z; z) = k$ implies $z \in \text{bd } Q$ and $z \in \text{bd lev}_k f$. Since $Q \supset \text{lev}_k f$ and both sets are convex, there must exist a hyperplane through z, supporting both Q and $\text{lev}_k f$; furthermore one also has $z \in \text{bd } P(k)$. Thus, since $P(k) \subset Q$, the set $P(k)$ has a common boundary with the set $\text{lev}_k f$ within the set X. Convexity implies $P(k) \cap X = \text{lev}_k f \cap X$. Q.E.D.

COROLLARY L8.1: Consider the *maximal* extension L_k of $\text{lev}_k f$ with respect to X; then one has $\text{lev}_k f \subset P(k) \subset L_k$ and $(P(k) \cap \text{Int } X) = (L_k \cap \text{Int } X) = (\text{lev}_k f \cap \text{Int } X)$.

DEFINITION 8: A *polaroid extension* PF of the function f with respect to the set X is a function whose level sets are

$$\text{lev}_k\, PF = P(k)$$

i.e.,

$$PF(y) = \inf\{+\infty, k | y \in P(k)\}.$$

Note that PF depends (implicitly) on the polarization φ of f and that there are, in general, several polaroid extensions of the same f. One may also observe that the definition of PF does not require f to be convex; PF is always quasi-convex however. Furthermore one has the following alternate definition of $PF(y)$ for $y \in$ dom PF (i.e., $PF(y) < +\infty$):

$$PF(y) = \inf\{k | \varphi(x; y) \leq k, \forall x \in X \cap \text{lev}_k f\}$$

$$= \sup\{k | \varphi(x; y) \geq k, \text{ some } x \in X\};$$

hence

$$PF(y) = \max_{x \in X} \varphi(x; y)$$

and the above expression can be used to compute any value of PF.

COROLLARY L8.2: $f(x) \leq PF(x) \leq MQCF(x)$, $\forall x$ and equality holds throughout $\forall x \in$ Int X.

PROOF: Follows from Corollary L8.1.

There exists no general dominance theorem between MCF and PF; however, one obviously has

COROLLARY L8.3: If the polaroid extension PF is convex then

$$PF(x) \geq MCF(x) \text{ (and } PF(x) = MCF(x) = f(x), \forall x \in \text{ Int } X).$$

PROOF: By maximality of MCF.

In conclusion, we note that polaroid extensions may constitute a viable alternative to the definitions of Sections 1 and 2, particularly when the non-linear programs

$$\text{maximize } \varphi(x; y), \text{ subject to } x \in X$$

are conveniently solved for each parameter y. Since it is necessary to produce the *global* optimum of the above program, one will naturally restrict the choice of φ to functions of the following type:

$\forall y$, $\varphi(x; y)$ *pseudo-concave* (in x); remember that we also assume

$\forall x$, $\varphi(x; y)$ *quasi-convex* (in y), in order to obtain convex polaroids.

LEMMA 9: Let φ satisfy the above conditions, and X be a convex set, then the program

$$\max_{x \in X} \varphi(x; y)$$

is convex $\forall y$, i.e., every local optimum is a global optimum.

PROOF: Omitted.

It seems natural, at this point, to question whether the abstract definition of a polaroid extension is really useful, and in particular if there are cases where a concrete polarization φ can be explicitly given; the following sections will now show that indeed one can easily find polarizations of a given function f which are both *strong* (in the sense of the Dominance Theorem T4.1) and *implementable*.

6. DIFFERENTIABLE f

In this section we present some results related to Tuy's note [5]. First we give by means of a polaroid extension an analytical characterization of Tuy's convex extension; a family of polaroid extensions is then constructed which satisfies a dominance property (Theorem 8). The practical significance of this theory is briefly illustrated with a quadratic example while a more comprehensive description of this special case can be found in [1] for the concave and [4] for the general case.

Define the polarization

$$\varphi(x, y)=f(x)+(y-x)\ \Delta f(x).$$

Clearly $\varphi(x, x)=f(x)$ and since φ is linear in y, Theorem 5 holds; moreover,

LEMMA 10: If f is *quasi-convex*, φ is *proper*.

PROOF: Quasi-convexity of f implies, in the case of differentiality:

$$\max\{f(x), f(y)\}\geq f(x)+(y-x)\nabla f(x);$$

hence $\varphi(x; y)\leq \max\{f(x), f(y)\}$. Q.E.D.

We now restrict ourselves to a convex function f and show that the corresponding *polaroid extension PF is the maximal convex extension MCF.* (Lemms 11 and 12).

LEMMA 11: *PF* is *convex* and $PF(y)\leq f(y)$ when f is pseudo-convex.

PROOF: We only need consider $y\in$ dom PF; take

$$y=\lambda y^1+(1-\lambda)y^2$$

$$PF(y)=\max_{x\in X}\varphi(x; y)=\max_{x\in X}[\lambda\varphi(x; y^1)+(1-\lambda)\varphi(x; y^2)]=f(\hat{x})+[\lambda y^1+(1-\lambda)y^2-\hat{x}]\nabla f(\hat{x})$$

$$=\lambda[f(\hat{x})+(y^1-\hat{x})\nabla f(\hat{x})]+(1-\lambda)[f(\hat{x})+(y^2-\hat{x})\nabla f(\hat{x})]\leq\lambda\max\varphi(x; y^1)+(1-\lambda)\max_{x\in X}\varphi(x; y^2)$$

$$=\lambda PF(y^1)+(1-\lambda)PF(y^2).$$

Also $PF(y)=\max_{x\in X}\varphi(x; y)=f(\hat{x})+(y-\hat{x})\nabla f(\hat{x})\leq f(y)$ by pseudo-convexity of f.

LEMMA 12: PF is *maximal.*

PROOF: We must show that for any convex function $g(x)$ with $g(x)=PF(x)=f(x)$, $\forall x\in X$ one has

$$g(y)\geq PF(y),\ \forall y\in\mathbb{R}^n.$$

Since g and f agree identically on X, g is subdifferentiable; furthermore one has

$$\nabla f(x)\in\partial g(x),\ \forall x\in X$$

where $\partial g(x)$ is the subdifferential of g at x. Now the subgradient inequality gives

$$g(y)\geq g(x)+(y-x)u,\ \forall u\in\partial g(x),\ \forall y\in\mathbb{R}^n;$$

in particular for $u=\nabla f(x)$.

Now given $y\in\mathbb{R}^n$, choose $\bar{x}\in X$ such that

$$PF(y)=\max_{x\in X}\varphi(x; y)=f(\bar{x})+(y-\bar{x})\nabla f(\bar{x});$$

one has

$$g(y) \geq g(\bar{x}) + (y-\bar{x})\nabla f(\bar{x}) = f(\bar{x}) + (y-\bar{x})\nabla f(\bar{x}) = PF(y) \qquad \text{Q.E.D.}$$

It is interesting to note the following additional property of PF: Take $y \in \mathbb{R}^n - X$; assume $x' \in X$ is such that

$$\varphi(x'; y) = f(x') + (y-x')\nabla f(x') = \max_{x \in X} \varphi(x; y) = PF(y).$$

Then

$$PF(z) = f(x') + \lambda(y-x')\nabla f(x'), \quad z = \lambda y + (1-\lambda)x', \quad \lambda \in [0.1]$$

because PF is convex and satisfies by construction

$$PF(z) \geq \varphi(x'; z) \text{ (since } x' \in X)$$

with equality holding at $z = x'$ and $z = y$; i.e., the convex extension PF is *linear* on the segment $x', y)$.

COROLLARY L12.1: $PF = MCF$.

Note that convexity of f is not essential here; however, if f were not convex one could have $PF(x) \neq f(x)$, for some $x \in X$, since PF is convex. It should also be noted that PF is only maximal with respect to other *convex* extensions; if quasi-convex extensions are taken into consideration, maximality is, in general, lost, as shown by Theorem 8 below.

The polaroid extension PF can be extended to the case where the function f is only required to be convex and continuous on X (i.e., only subdifferentiable instead of differentiable); this merely involves replacing the gradient by a subdifferential.

It is interesting to note that although PF is the maximal convex extension of a convex (differentiable) function f, it is really not the best extension which can be obtained by this type of polarization; to illustrate this point, we now imbed PF in a family of polaroid extensions PF_α (with $PF_1 = PF$)

$$\varphi_\alpha(x; y) = f(x) + \alpha(y-x)\nabla f(x), \quad \alpha > 0.$$

and prove the following dominance theorem.

THEOREM 8: $PF_{\alpha_1}(x) \geq PF_{\alpha_2}(x)$ if $\alpha_1 \geq \alpha_2$.

PROOF: We show that the corresponding polaroids satisfy $P_{\alpha_1}(k) \subset P_{\alpha_2}(k)$. Take $y_1 \notin \text{bd } P_{\alpha_1}(k)$, and choose $x_1 \in X \cap \text{lev}_k f$ such that

$$\varphi_{\alpha_1}(x_1; y_1) = k.$$

Suppose $y_1 \in P_{\alpha_2}(k)$; then there must exist a point $x_2 \in X \cap \text{lev}_k f$ such that

$$\varphi_{\alpha_2}(x_2; y_1) = k' > k.$$

It follows that $\alpha_2(y_1 - x_2)\nabla f(x_2) = k' - f(x_2) > 0$ (since $f(x_2) \leq k$) and hence

$$(y_1 - x_2)\nabla f(x_2) > 0.$$

Furthermore, $y_1 \in P_{\alpha_1}(k)$ implies

$$f(x_2) + \alpha_1(y_1 - x_2)\nabla f(x_2) = f(x_2) + \alpha_2(y_1 - x_2)\nabla f(x_2) + (\alpha_1 - \alpha_2)(y_1 - x_2)\nabla f(x_2) \leq k$$

hence

$$(\alpha_1-\alpha_2)(y_1-x_2)\nabla f(x_2)\leq k-k'<0$$

which implies $(\alpha_1-\alpha_2)<0$.

A contradiction the hypothesis $\alpha_1\geq\alpha_2$; thus, $y_1\in P_{\alpha_2}(k)$. Q.E.D.

An Illustration: The Quadratic Case [1]

Consider

$$f(x)=cx+\frac{1}{2}xCx,$$

where f is convex, i.e., C positive semi-definite. The maximal convex extension of f with respect to a feasible set X is defined by the polarization (see L12.1)

$$\varphi(x;y)=cx+\frac{1}{2}xCx+cy+yCx-cx-xCx=-f(x)+[cx+cy+yCx]$$

COROLLARY L12.2: The maximal convex extension of a *quadratic* (convex) function can be obtained by solving *convex programs* of the type

$$\max_{x\in X}\varphi(x;y),$$

where y is given and $\varphi(x;y)$ is a *concave* function (of x).

PROOF: Since f is convex, $-f$ is concave and $\varphi=-f+$linear terms is also concave; hence, the program is convex. Q.E.D.

But the family of polarizations $\varphi_\alpha(x;y)=f(x)+\alpha(y-x)\nabla f(x)$, $(\alpha>0)$ obeys the dominance relation of Theorem 8. Moreover, for the quadratic (convex) case, observe that

$$\varphi_{1/2}(x;y)=\frac{1}{2}[cx+cy+yCx]$$

is *bilinear*. Thus, in this case the *convex* programs of Corollary L12.2 are *linear* programs. Since one has

$$\varphi_\alpha(x;y)=-(2\alpha-1)f(x)+\alpha[cx+cy+yCx],$$

it is apparent that for any given y the function $\varphi_\alpha(x;y)$ is *convex* in x when $0<\alpha<1/2$; thus, a polaroid extension of this type "better" than $PF_{1/2}$ (in the sense of Theorem 8) can only be obtained at the cost of solving a (non-convex) *concave quadratic program* in Corollary L12.2; this is of course of little practical value since such programs are often of the same degree of difficulty as the original concave program: $\max f(x)$, subject to $x\in X$. However, in certain cases, where the polyhedral set used to define the polaroid has a simple structure (for instance, the simplical set S of Section 7) this can be efficient (see [4]).

7. POLAROIDS AND CUTTING PLANES

As mentioned in Section 6, one of the applications of polaroid extensions is (quasi-) concave programming. We now indicate how cutting planes can be constructed from a given extension. In the second part of this section we present a new (interactive) construction of the "deepest" cut to be generated in the present framework.

7.1. Cutting Planes

The inclusion test (see Section 0) of two convex sets can often be accomplished efficiently by a cutting plane algorithm, particularly when the feasible set X is polyhedral. (There exist other types of procedures, which are also based on convex extensions and can be used in this context (see, for instance [1])).

Take the current best value $\bar{k}=f(\bar{x})$ where $\bar{x}\in X$ is the corresponding solution; by hypothesis one has $\bar{x}\in \text{lev}_{\bar{k}}f$ and a cutting plane is constructed to discard a (cut-off) portion $S\cap X$ which cannot contain a point y better than $\bar{x}$, i.e., $S\cap X\subset \text{Int}\,(\text{lev}_{\bar{k}}f)$, or $f(y)<\bar{k}$, $\forall y\in S\cap X$ where the set S is used to denote a cut (see [2]). (Here f should be assumed strictly quasi-convex and lower semi-continuous). Cutting planes are a convenient way to define such a cut-set S which is then a half-space; thus, in essence cuts are used to reduce the feasible set X and thereby facilitate the inclusion test.

When X is polyhedral, one can define a simplicial cut-set S with $\bar{x}$ a vertex of X. Let $\bar{x}$ be obtained from an L.P. tableau of the form

$$x_i=\bar{x}_i-\sum_{j\in\bar{N}} t_j\bar{a}_{ij},\ i\in M$$

where $\bar{N}$ is the non-basic index set corresponding to $\bar{x}$. The cut-simplex S can be determined as the convex hull of the $(n+1)$ points (see [3, 5]) $\{\bar{x}, \bar{u}_j, j\in\bar{N}\}$ where $\bar{u}_j$ is the intersection point of the ray $u_j(t_j)=\bar{x}-t_j\bar{a}_j$, $t_j\geq 0$ with the boundary of the set $\text{lev}_{\bar{k}}f$ (or (better) $\text{lev}_{\bar{k}}MCF$, $PF(\bar{k})$, or $\text{lev}_{\bar{k}}MQCF$) i.e., $\bar{u}_j=u_j(\bar{t}_j)$, $\bar{t}_j\geq 0$. Then one has

$$S=\left\{x=x(t)\,\middle|\,\sum_{j\in\bar{N}}\frac{1}{\bar{t}_j}t_j<1,\ t_j\geq 0,\ \forall j\in\bar{N}\right\}.$$

COROLLARY T4.2: The cut-simplices generated by intersection with convex extensions satisfy the following *uniform dominance* relation: $S_f\subset S_{MCF}\subset S_{MQCF}$

7.2. Polaroids

The main purpose of a polaroid extension in the construction of a cut is to insure depth and validity; and the following immediate observations come to mind:

a) Validity is defined by $x\in S\cap X\Rightarrow x\in \text{lev}_{\bar{k}}f$ i.e., the points of X *which are actually cut-off* must lie in $\text{lev}_{\bar{k}}f$: $(S\cap X)\subset \text{lev}_{\bar{k}}f$.

b) Polaroids P^* of a given set P can be viewed here as a tool to generate cut-sets $S\in P^*$ because under proper assumptions on f, one has $P^*\cap P=P\cap \text{lev}_k f$ (Lemma 8 and corollaries).

c) Finally, polaroids enjoy the following inclusion reversing property, (see [3]).

$$Q\subset P\Rightarrow P^*\subset Q^*$$

Thus we may now combine a), b) and c) and define the following:

1) Choose the set $Q=(X\cap S\cap \text{lev}_k f)$ where S is a cut-simplex (defined under 3) below).

2) Define the corresponding polaroid

$$CP(k)=\{y|\varphi(x;y)\leq k,\ \forall x\in Q\}$$

$$=\{y|\varphi(x;y)\leq k,\ \forall x\in(X\cap S\cap \text{lev}_k f\}\supset\{y|\varphi(x;y),\ \forall x\in(X\cap \text{lev}_k f)\}=P(k)\supset \text{lev}_k f.$$

(See Lemma 8)

3) S is the cut-simplex obtained by interesction with the polaroid CP; note that Q is defined by S so that S really plays the role of a *variable parameter* in this construction.

Now, by virtue of the inclusion reversing property c), one may define the following (dynamic) geometrical representation of the interaction between S and the polaroid CP:

- Starting with a small S (for instance $X \cap S = \{\bar{x}\}$). one obtains a very large CP (since $\{\bar{x}\}^*(k)$ is a halfspace);
- Since CP is "enormous", a larger cut simplex $\tilde{S} \supset S$ can be generated by ray intersection with CP.
- But, now the corresponding $\widetilde{CP}$ is shrinking ($\widetilde{CP} \subset CP$) because of the inclusion reversing property c).

Thus, one easily imagines that at some point, this mechanism will stop in some equilibrium position (usually not unique) where S can no longer be made larger without violating the condition $S \cap X \subset CP$ required to guarantee validity.

The set $CP(k)$ is called *cutting polaroid* to emphasize the interaction between the cut and polaroid constructions: for the vertices of the cut-simplex S one has $\bar{u}_j \in \text{bd } CP(k)$, $\forall j \in \bar{N}$, that is:

$$\max \varphi(\bar{u}_j, x) = k$$

$$\text{subject to } x = \bar{x} - \sum_{j \in \bar{N}} \bar{a}_j t_j \in X$$

$$\sum_{j \in \bar{N}} \frac{1}{\bar{t}_j} t_j < 1,\ t_j \geq 0, \forall j;$$

where the $\bar{t}_j$ are parameters.

A computationally simpler variant of this scheme is to start with a given hyperplane supporting X at $\bar{x}$, say

$$\sum_{j \in \bar{N}} \alpha_j t_j = 0,\ \alpha_j > 0,\ \forall j \in \bar{N}$$

and to "push" it into the feasible set as far as possible to create the cut

$$\sum_{j \in \bar{N}} \alpha_j t_j \geq \delta (> 0)$$

and the corresponding cut-simplex S; the *cutting polaroid approach* is then used to determine the (largest) value $\delta > 0$ such that $\forall_j \in \bar{N}$:

$$\max \varphi(\tilde{u}_j, x) \leq k$$

$$\text{subject to } x \in X$$

$$* \qquad \sum_{j \in \bar{N}} \alpha_j t_j < \delta,\ t_j \geq 0,\ \forall j \in \bar{N}$$

with

$$\tilde{u}_j = \bar{x} - \bar{a}_j \left(\frac{\delta}{\alpha_j} \right).$$

In this case, there is only one parameter δ; note that δ merely corresponds to the slack variable of the cut constraint (*); otherwise δ only appears in the objective function φ.

Thus, the cutting polaroid approach provides a means to determine a deep valid cut by parametric optimization (with parameter δ). These programs have been studied in [4 and 8] in the quadratic case; this approach is more efficient than the (conventional) intersection approach proposed in [5], because this parametric program (*linear* in the quadratic case) is easily solved.

8. DISCRETE EXTENSIONS WITH RESPECT TO POLYHEDRAL SETS

When X is polyhedral, it is well known that the program

$\max f(x)$, subject to $x \in X$

(f convex)

is either unbounded or it possesses an *optimal vertex*; thus, the range of f may be restricted to the *finite* set of values $f(\text{vert } X)$ without essentially modifying the final outcome of the above optimization problem. A polaroid extension SPF can thus be defined as a *quasi-convex step function* of the following type:

$$\forall k \in f(\text{vert } X),\ \text{lev}_k\, PSF = \{y \mid \varphi(x; y) \leq k,\ \forall x \in \text{vert } X \text{ with } f(x) \leq k\};$$

for intermediate values $k \in (k_1, k_2)$ one has: $\text{lev}_k\, PSF = \text{lev}_{k_2}\, PSF$ where

$$k_1 \text{ and } k_2 \in f\ (\text{vert } X),\ k \notin f\ (\text{vert } X).$$

In fact, we can modify the definition of PSF further, by selecting in the definition of the polaroid only *local optima* of the initial non-linear problem, i.e., only the candidates for the global optimum, or (assuming differentiality):

$\text{lev}_k\, PSF = \{y \mid \varphi(x; y) \leq k$, for all those $x \in X \cap \text{lev}_k f$ satisfying the necessary conditions

$$(\xi - x)\nabla f(x) \leq 0,\ \forall \xi \notin X\}.$$

It is easily seen from the definition that $\text{lev}_k\, PSF \supset P(k) = \text{lev}_k\, PF$; i.e., one has the dominance relation $PSF(x) \leq PF(x)$; note also that for $x \in X - \{\text{vert } X\}$ one may have $PSF(x) < f(x)$ so that, strictly speaking, PSF is no longer an extension as defined in Section 1. However, PSF enjoys the same properties as another polaroid extension, if one replaces X by vert X; and we therefore do not duplicate the results of the preceding sections.

REMARKS:

a) The construction of PSF, in principle, requires one to solve (globally) an *extreme point program* of the type

$$\max_{x \in Q_0} \varphi(x; y),\ y \text{ given},\ Q_0 = (\text{vert } X \cap \text{lev}_k f).$$

This may seem a tremendous task, but any set Q satisfying $(X \cap \text{lev}_k f) \supset Q \supset Q_0$ may be used to deliver a bound k' which furnishes an approximation (from above) to PSF:

$$\max_{x \in Q_0} \varphi \leq \max_{x \in Q} \varphi = k'.$$

The step function PSF represents an improvement over the corresponding (i.e., with same polarization φ) polaroid extension PF; and this happens in two ways which are best illustrated by an application to concave programming.

(i) Because the parameter k can always be chosen to belong to $\Delta=\{k=f(x)|x\in \text{vert } X, x \text{ local optimum}\}$ the value of k is increased step-wise by discrete amounts, until the value

$$\bar{k}=\max_{x\in X} f(x)$$

is reached. In a practical sense, this can be realized during the construction of a cut-simplex S, for instance; one sets

$$k_{\text{new}}=\max\{k_{\text{old}}, f(\bar{x}), f(x^j), \forall j\in \bar{N}\}$$

where x^j are the neighboring (on X) vertices of $\bar{x}$ i.e.,

$$x^j=\bar{x}-\hat{t}_j\bar{a}_j$$

where $\hat{t}_j$ is the largest value $t_j\geq 0$ such that $x^j\in X$.

(ii) One may replace X by the set $\Gamma=\{x\in \text{vert } X, x \text{ locally optimal}\}$ in the construction of a *cutting polaroid* (see Section 7) and obtain $CP(k)=(\Gamma\cap S)^*(k)$.

ACKNOWLEDGMENT

I wish to express my thanks to R. Breu, who contributed to many improvements and corrections of the manuscript.

BIBLIOGRAPHY

[1] Balas, E. and C.-A. Burdet: "Concave Quadratic Programming," Management Science Research Report #299, Carnegie-Mellon University (1972).

[2] Burdet, C.-A., "Polaroids: A New Tool in Non-Convex and in Integer Programming," Naval Research Logistic Quarterly 20, 13–24 (1973).

[3] Burdet, C.-A., "On Polaroid Intersections," *Mathematical Programming in Theory and Practice*, P. Hammer and G. Zoutendijk eds., pp. 365–387 (North Holland, 1974).

[4] Burdet, C.-A., "On Linearly Constrained Non-Convex Quadratic Programs," W.P. 91–72–3, Graduate School of Industrial Administration, Carnegie-Mellon University (1972).

[5] Tuy, Hoang, "Concave Programming Under Linear Constraints," (Russian). *Doklady Akademii Nauk SSSR*, 1964. English translation in Soviet Mathematics, 1437–1440 (1964).

[6] Minkowski, H., "Theorie der Konvexen Körper insbesondere Begründung ihres Oberflächenbegriffes." Gesammelte Abhandlungen, 2 (Leipzig, 1911).

[7] Rockafellar, R. T.: *Convex Analysis*. (Princeton University Press, 1970).

A CUTTING PLANE ALGORITHM FOR THE BILINEAR PROGRAMMING PROBLEM

H. Vaish
California State University, Northridge
Northridge, California

C. M. Shetty
Georgia Institute of Technology
Atlanta, Georgia

ABSTRACT

In this paper we discuss the properties of a Bilinear Programming problem, and develop a convergent cutting plane algorithm. The cuts involve only a subset of the variables and preserve the special structure of the constraints involving the remaining variables. The cuts are deeper than other similar cuts.

I. INTRODUCTION

The Bilinear Programming Problem considered in this paper can be stated as:

BLP: Minimize $\phi(x, y)=c^t x+d^t y+x^t Cy$

(1) Subject to: $x\epsilon X_0=\{x\epsilon R^m | Ex\leq e,\ x\geq 0\}$

$y\epsilon Y_0=\{y\epsilon R^n | Fy\leq f,\ y\geq 0\}$

Without loss of generality, we will assume that X_0 and Y_0 are bounded. In spite of its special structure, problem BLP is a mathematical statement of a number of practical problems, for example, location-allocation problems, orthogonal production scheduling, multi-stage assignment problem, etc. (See [23] and [28]).

An important property of BLP to observe is that even though ϕ can be shown to be not quasi-concave, the optimal (x^*, y^*) is attained at an extreme point of $X_0\times Y_0$, i.e., x^* will be an extreme point of X_0 and y^* is an extreme point of Y_0 [23, 29]. It seems reasonable that in solving BLP this extreme point optimality property of BLP should be taken advantage of. However, ϕ is not explicitly quasi-convex so that local minima can and do exist [28]. This is the essential difficulty in solving BLP.

Problem BLP can be regarded as a Quadratic Programming problem in which the objective function need not be convex. A number of algorithms have been proposed for solving this class of problems. One group of these algorithms generates a sequence of expanding polytopes such that the minimum over each is known. This annexation procedure terminates when some polytope in the sequence contains the original polytope [13, 27, 29]. Some alternative approaches are discussed in [10, 24].

The strategy we will use to solve the problem will be to develop a series of cuts such that no point with a lower value of $\phi(x, y)$ than the current best available is deleted. The approach is therefore the *piecewise-strategy* discussed by Geoffrion [14] for solving large-scale problems. The process is repeated until all the feasible region has been explored.

II. CUTTING PLANE STRATEGIES

Cutting planes have been used in integer programming for some time. Recent developments in [1, 2, 6, 15, 17, 19, 30] show how valid cuts can be generated using certain convex sets. On the other hand, several authors have used the cutting plane approach to solve nonconvex problems, e.g., see [3, 5, 11, 12, 13, 21, 26, 27]. The recent work of Burdet unifies these approaches through the use of Polaroids [6, 7, 9] and the related more general concept of level sets of convex gauge functions [8]. Our main concern here is to use Burdet's approach to solve problem *BLP*. However, as we will see, some modifications need to be done.

Consider a problem P: Min. $f(x)$, subject to $x \epsilon S$, where S is a polyhedral (compact) set and f is nonconvex. A *local star minimum* of P is an extreme point $\bar{x}$ such that $f(\bar{x}) \leq f(x)$ for each $x \epsilon N(\bar{x})$ where $N(\bar{x})$ denotes the adjacent extreme points to $\bar{x}$. A *local minimum* of P is a point $\bar{x}$ such that $f(\bar{x}) \leq f(x)$ for each $x \epsilon B_\delta(\bar{x}) \cap S$, where $B_\delta(\bar{x})$ is a δ-neighborhood around $\bar{x}$. If f is quasi-concave, then a local star minimum is also a local minimum. In such a case, a cutting plane can be developed from a local star minimum $\bar{x}$ which cuts off $\bar{x}$ but not any other improving point as was done in [5, 27]. On the other hand, if the assumption does not hold, a local minimum $\bar{x}$ is needed to define a cutting plane [3, 11, 26].

There is yet another important property we would like to preserve. Cuts involving variables associated with both X_0 and Y_0 sets will destroy their special structure. There are problems wherein one of the sets does have a special structure for which efficient algorithms exist that can be used to solve sub-problems in the solution procedure. In the location-allocation problem, the Y_0 set can be made to represent the transportation problem constraints.

Konno [23] has discussed this strategy in detail for the *BLP*. At a local star optimim $(\bar{x}, \bar{y})$ he defines a vector g^t and finds the parameter value $\bar{\sigma}$ such that the cut $g^t x \geq \bar{\sigma}$ deletes $\bar{x}$ but no point x such that $\phi(x, y) \leq k - \epsilon$ where k is the objective function value for the current best solution and $\epsilon > 0$ is a predetermined value. The procedure thus yields an *ϵ-optimal solution*, i.e., a point (x^*, y^*) such that $\phi(x^*, y^*) \leq \phi(x, y) + \epsilon$ for all $x \epsilon X_0$ and $y \epsilon Y_0$.

Gallo and Ülkücü [13] have developed a cutting plane algorithm for *BLP* similar to the one discussed in this paper, where the cuts are applied in the X-set. Using duality theory they consider the following equivalent problem: Min. $\psi(x, u) = (c^t x + \text{Max } f^t u)$ subject to $x \epsilon X_0$ and $u \epsilon U = \{u: F^t u \leq d + Cx, u \leq 0\}$. For any $x^k \epsilon X_0$, let u^k be the point obtained by solving Max $f^t u$ subject to $F^t u \leq d + Cx^k$, $u \leq 0$. A cutting plane is generated from a vertex x^k such that $\psi(x^k, u^k) > k$ where k is the current best value of the objective function. Thus, if the vertex x^k is in fact the current best, one has to move to the point yielding a poorer value of Ψ. If no such point exists, special steps need be taken. The algorithm is not guaranteed to converge.

Earlier we have indicated that cuts involving both the x and y variables can be generated from a local minimum. However, if we want to develop a cutting plane involving only the x-variables and yet be convergent, we need to develop the cut at $(\bar{x}, \bar{y})$ which is more than a local optimum.

Such a point which is adequate is defined below. Throughout the rest of the paper by X_0 we mean the original feasible region X_0 or its subset obtained after the introduction of cuts.

DEFINITION: An extreme point $(\bar{x}, \bar{y})$ is called a *pseudo-global minimum if* $\phi(\bar{x}, \bar{y}) \leq \phi(x, y)$ for each $x \epsilon B_\delta(x) \cap X_0$ and for each $y \epsilon Y_0$.

Note that $(\bar{x}, \bar{y})$ is an extreme point of the constraints of *BLP* if and only if $\bar{x}$ is an extreme point of X_0 and $\bar{y}$ is an extreme point of Y_0. Further, an extreme point is adjacent to $(\bar{x}, \bar{y})$ if and only if it is of the form $(x^i, \bar{y})$ or $(\bar{x}, y^i)$ where $x^i \epsilon N(\bar{x})$ and $y^i \epsilon N(\bar{y})$. We will now characterize the various forms of "optimality" for *BLP* which will be used later in the development of the algorithms. Consider an extreme point $(\bar{x}, \bar{y})$. One can readily transform the origin to this point. Konno [23] has shown that such a transformation will yield a problem of the form given in (1) (with the additional property that $e \geq 0, f \geq 0$). The following theorems refer to the transplated problem with

$$(\bar{x}, \bar{y}) = (0, 0).$$

THEOREM 1: The origin (0, 0) is a local star minimum of *BLP* if and only if $x=0$ solves the linear program $P1$: Min $\phi(x, 0)$, $x \epsilon X_0$, and $y=0$ solves the linear program $P2$: Min $\phi(0, y)$, $y \epsilon Y_0$.

PROOF: Let $x=0$ solve $P1$ and $y=0$ solve $P2$. Then $\phi(0, 0) \leq \phi(x, 0)$, $x \epsilon X_0$. In particular, $\phi(0, 0) \leq \phi(x^i, 0)$, $x^i \epsilon N(0)$. Similarly, $\phi(0, 0) \leq \phi(0, y^i)$, $y^i \epsilon N(0)$. Hence,

$$\phi(0, 0) \leq \phi(x^i, y^i), \qquad (x^i, y^i) \epsilon N(0, 0).$$

Hence, (0, 0) is a local star minimum. Conversely, let (0, 0) be a local star minimum. Hence. $\phi(0, 0) \leq \phi(0, y^i)$, $y^i \epsilon N(0)$. Consider problem $P2$. Suppose we apply the simplex algorithm and obtain a basic feasible solution corresponding to $y=0$. Since $\phi(0, 0) \leq \phi(0, y^i)$ for each $y^i \epsilon N(0)$, the simplex method will terminate and 0 is a solution to $P2$. Similarly, $x=0$ solves $P1$.

THEOREM 2: The origin (0, 0) is a global minimum of *BLP* if $c^t x \geq 0$, $d^t y \geq 0$ and $x^t C y \geq 0$ for each $x \epsilon X_0$, $y \epsilon Y_0$.

PROOF: Trivial, since the hypothesis implies $\phi(x, y) \geq \phi(0, 0)$ for each $x \epsilon X_0$ and $y \epsilon Y_0$.

THEOREM 3: The origin (0, 0) is a local minimum of *BLP* if and only if for each $x \epsilon X_0$ and $y \epsilon Y_0$

(i) $c^t x \geq 0$ and $d^t y \geq 0$, and

(ii) if $x^t C y < 0$ then $(c^t x + d^t y) > 0$.

PROOF: Suppose conditions (i) and (ii) hold. Then for $\epsilon > 0$ small enough $\phi(\epsilon x, \epsilon y) = \epsilon(c^t x + d^t y + \epsilon x^t C y) \geq 0 = \phi(0, 0)$ since the term within parenthesis is positive. Hence (0, 0) is a local minimum.

Now suppose (0, 0) is a local minimum. If $c^t x < 0$, then $\phi(\epsilon x, 0) = \epsilon c^t x < 0 = \phi(0, 0)$ contradicting the local optimality of (0, 0). Hence, $c^t x \geq 0$, and likewise $d^t y \geq 0$. Now if $x^t C y < 0$ and $(c^t x + d^t y) \leq 0$, then $\phi(\epsilon x, \epsilon y) = \epsilon(c^t x + d^t y + \epsilon x^t C y) < 0 = \phi(0, 0)$ for all $\epsilon > 0$ again contradicting the local optimality of (0, 0). Hence, both conditions (i) and (ii) hold.

THEOREM 4: The origin (0, 0) is a local star minimum of *BLP* if and only if $c^t x \geq 0$ and $d^t y \geq 0$ for $x \epsilon X_0$, $y \epsilon Y_0$.

PROOF: Suppose (0, 0) is a local star minimum. Then from Theorem 1 $x=0$ is a solution to the problem:

$$\text{Min } \phi(x, 0) = c^t x, \quad x \epsilon X_0.$$

Hence, $c^t x \geq 0$ for each $x \epsilon X_0$. Likewise $d^t y \geq 0$ for each $y \epsilon Y_0$. Conversely, let $c^t x \geq 0$ and $d^t y \geq 0$ for each $x \epsilon X_0$ and $y \epsilon Y_0$. But $\phi(x, 0) = c^t x$ and $\phi(0, 0) = 0$. Hence $x = 0$ solves: Min $\phi(x, 0)$, $x \epsilon X_0$. Likewise $y = 0$ solves: Min $\phi(0, y)$, $y \epsilon Y_0$. Then from Theorem 1 we have (0, 0) is a local star minimum.

THEOREM 5: The origin (0, 0) is a *pseudo-global minimum* of *BLP* if and only if for each $x \epsilon X_0$ and $y \epsilon Y_0$.

(i) $d^t y \geq 0$, and

(ii) if $d^t y = 0$ then $(c^t x + x^t C y) \geq 0$.

PROOF: Suppose (0, 0) is a pseudo-global minimum. By definition it readily follows that it is a local minimum. Hence, from Theorem 3 we have $d^t y \geq 0$. Now for $x \epsilon X_0$, $\alpha x \epsilon B_\delta(0) \cap X_0$ for $\alpha > 0$ small enough. Then $\phi(\alpha x, y) = \alpha(c^t x + x^t C y) + d^t y \geq \phi(0, 0) = 0$ for each $y \epsilon Y_0$ since (0, 0) is a pseudo-global minimum. Hence, if $d^t y = 0$ then $c^t x + x^t C y \geq 0$.

Now assume that conditions (i) and (ii) hold. Consider the m edges from $x = 0$ to the adjacent extreme points. Let $0 \neq x^i \epsilon X_0$ be on the i^{th} edge. Let $y^1, \ldots, y^K$ denote the K extreme points of Y_0 and let

$$\alpha_{ik} = \begin{cases} -d^t y^k / (c^t x^i + x^{it} C y^k) & \text{if } (c^t x^i + x^{it} C y^k) < 0 \\ \infty & \text{otherwise} \end{cases}$$

Note that $\alpha_{ik} > 0$ by assumption (i) and (ii). Let

$$\alpha_i = \min_{k=1, \ldots, K} \alpha_{ik}.$$

If all $\alpha_{ik} = \infty$ we will let α_i to be an arbitrarily large number.
Then $\phi(\alpha_i x^i, y^k) = \alpha_i(c^t x^i + x^{it} C y^k) + d y^k \geq 0$ from conditions (i) and (ii) and the definition of α_i. But, the minimum of $\phi(x, y)$ for a fixed x is achieved at an extreme point of Y_0. Hence,

$$\operatorname{Min}_{y \epsilon Y_0} \phi(\alpha_i x^i, y) \geq 0$$

for each $i = 1, \ldots, m$.

Now let $S = \text{Conv.}\ [0, \alpha_i x^i \text{ for } i = 1, \ldots, m]$. Then Min $\phi(x, y)$, $x \epsilon S$, $y \epsilon Y_0$ is achieved at an extreme point of S, which is 0 or of the form $\alpha_i x^i$. Hence,

$$\operatorname{Min}_{\substack{x \epsilon S \\ y \epsilon Y_0}} \phi(x, y) \geq \text{Min. } \{\phi(0, 0), \operatorname{Min}_{y \epsilon Y_0} \phi(\alpha_i x^i, y)\} \geq 0.$$

Now select a $\delta > 0$ such that $B_\delta(0) \cap X_0 \subset S$. Then Min $\phi(x, y) \geq 0$ for each $x \epsilon B_\delta(0) \cap X_0$ and for each $y \epsilon Y_0$. That is, (0, 0) is a pseudo-global minimum.

It may be mentioned that the corresponding theorems in [23] can be obtained as special cases of Theorems 2, 3, and 4. Further, from Theorems 3 and 4 we may observe that a local minimum is always a local star minimum. The converse may not be true as shown in [28]. Further, from the definitions one can note that a pseudo-global minimum is always a local minimum.

III. GENERATION OF A PSEUDO-GLOBAL MINIMUM

Before discussing the procedure for getting a pseudo-global minimum, we will review the method due to Balas [1] for identifying precisely m edges incident on an extreme point $\bar{x}$, which also leads to the resolution of degeneracy.

Let $\bar{x}$ be an extreme point of X_0 and let p_j, $j\epsilon J$, be the m nonbasic variables at $\bar{x}$, where J is the index set for the nonbasic variables. Denoting by $\bar{e}^j$ the columns of the simplex tableau in Tucker form (extended form), the m-vector x can be written as:

$$x=\bar{x}-\sum_{j\epsilon J}\bar{e}^j p_j$$

which satisfies the constraint $Ex\leq e$ (but not necessarily the nonnegativity restriction) for any $p_j\geq 0$. Let M and N denote the index sets associated with the constraints $Ex+u=e$ and $x\geq 0$ respectively, i.e.,

$$X_0=\left\{x\epsilon R^m;\ \sum_{j=1}^{m} e_{ij}x_j+u_i=e_i,\ i\epsilon N \text{ and } x_j\geq 0,\ j\epsilon M\right\}.$$

Given a basic feasible solution $(\bar{x}, \bar{u})$, let $N^0=\{i\epsilon N: u_i \text{ is basic and } \bar{u}_i=0\}$, $M^0=\{j\epsilon M: x_j \text{ is basic and } \bar{x}_j=0\}$. Let

$$X'_0=\left\{x\epsilon R^m:\ \sum_{j=1}^{m} e_{ij}x_j+u_i=e_i,\ i\epsilon N-N^0;\ x_j\geq 0,\ j\epsilon M-M^0\right\}.$$

Then clearly, $X_0\subset X'_0$ since X'_0 is obtained by deleting constraints of X_0.

THEOREM 6 [1]: Let $\bar{x}$ be an extreme point solution obtained by solving a linear program: Min. $\beta' x$, $x\epsilon X_0$. Let X'_0 be defined as above. Then $\bar{x}$ is a vertex of X'_0 and $\beta' x=\beta'\bar{x}$ is a supporting hyperplane for X'_0. Besides, X'_0 has precisely m distinct edges incident on $\bar{x}$ and each half line

$$\xi^j=\{x:\ x=\bar{x}-\bar{e}^j\lambda_j,\ \lambda_j\geq 0\},\ j\epsilon J$$

contains exactly one such edge.

We will now present two algorithms to generate a pseudo-global minimum.

ALGORITHM A.

1. Find a feasible extreme point x^1 of X_0.
2. (a) Solve: Min $\phi(x^1, y)$, $y\epsilon Y_0$, to yield an optimal y^1.
   (b) Solve: Min $\phi(x, y^1)$, $x\epsilon X_0$, to yield an optimum x^2.

Repeat until the procedure converges to a point $(\bar{x}, \bar{y})$, which clearly is a local star minimum.

3. Generate all alternative optimal extreme point solutions $y^1, \ldots, y^k (k\geq 1)$ to Min. $\phi(\bar{x}, y)$, $y\epsilon Y_0$. Solve: Min $\phi(x, y^i)$, $x\epsilon X_0$ for $i=1, \ldots, k$ to yield solutions $x^1, \ldots, x^k$. If $\phi(\bar{x}, \bar{y})\leq\phi(x^i, y^i)$ for all i, terminate; $(\bar{x}, \bar{y})$ is a pseudo-global minimum. If $\phi(x^r, y^r)<\phi(\bar{x}, \bar{y})$ for some r go to step 2(a) with $x^1=x^r$.

ALGORITHM B.

1 and 2. Find a local star minimum as in steps 1 and 2 of Algorithm A.

3. Let $x^1, \ldots, x^m$ be the adjacent extreme points of $\bar{x}$. Solve: Min $\phi(x^i, y)$, $y\epsilon Y_0$, to yield solutions $y^1, \ldots, y^m$. If $\phi(\bar{x}, \bar{y})\leq\phi(x^i, y^i)$ for all i, terminate; $(\bar{x}, \bar{y})$ is a pseudo-global minimum. If $\phi(x^r, y^r)<\phi(\bar{x}, \bar{y})$ go to step 2(b) with $y^1=y^r$.

It is intuitively evident that Algorithm B yields a pseudo-global minimum. However, to implement it we need a ready means of identifying the adjacent extreme points of $\bar{x}$. Also note that both of the algorithms are finite.

LEMMA 7: Let the origin be translated to the point $(\bar{x}, \bar{y})$ obtained either by Algorithm A or by Algorithm B. Then (0, 0) is a pseudo-global minimum.

PROOF: First consider Algorithm A. The algorithm is terminated at a point $(\bar{x}, \bar{y})$ which is clearly a local star optimum from Theorem 1. Hence, from Theorem 4 we have $d^t y \geq 0$. Now consider a $\hat{y} \epsilon Y_0$ satisfying $d^t \hat{y} = 0$. Clearly it is an alternative optimum to Min $\phi(0, y)$, $y \epsilon Y_0$ and can be expressed as

$$y = \sum_{i=1}^{k} \lambda_i y^i, \ \sum \lambda_i = 1, \ \lambda_i \geq 0$$

where y^i are the alternative extreme point optima at step 3. Hence, for $x \epsilon X_0$ using the notation of step 3,

$$c^t x + x^t C\hat{y} = \phi(x, \hat{y}) = \phi(x, \sum \lambda_i y^i) = \sum \lambda_i \phi(x, y^i) \geq \sum \lambda_i \phi(x^i, y^i) \geq \phi(0, 0) = 0.$$

Hence, from Theorem 5, (0, 0) is a pseudo-global minimum.

Now consider Algorithm B. Let $S = \text{conv } [0, x^1, \ldots, x^m]$ where $x^i \epsilon N(0)$. Then Min. $\phi(x, y)$, $x \epsilon S$, $y \epsilon Y_0$ is achieved at an extreme point of S and Y_0. From step 3 of the algorithm

$$\phi(0, 0) \leq \phi(x^t, y^t) \leq \phi(x, y)$$

for each $x \epsilon S$, $y \epsilon Y_0$. Select a $\delta > 0$ such that $B_\delta(0) \cap X_0 \subset S$. Then clearly (0, 0) is a pseudo-global minimum by definition.

IV. A CONVERGENT CUTTING PLANE ALGORITHM

We will now show how a cut can be generated from a pseudo-global minimum using the concept of generalized polars [6, 7].

DEFINITION: The *generalized polar* of Y_0 for a given scalar k is given by $Y^0(k) = \{x : \phi(x, y) \geq k \text{ for all } y \epsilon Y_0\}$

By definition $Y^0(k)$ contains no point $x \epsilon X_0$ such that $\phi(x, y) < k$ for some $y \epsilon Y_0$. Hence, if k is the current best value of the objective function, the problem is solved if $X_0 \subset Y^0(k)$.

Further, it can easily be verified that $Y^0('k)$ is compact and convex. As a matter of fact it is polyhedral since it can be shown [28] that

$$Y^0(k) = \{x; \ \phi(x, y^i) \geq k \text{ for all } y^i \epsilon V\}$$

where V is the set of extreme points of Y_0.

We will now discuss how a valid cut can be generated involving only the x-variables using generalized polars. Let $(\bar{x}, \bar{y})$ be a pseudo-global minimum of *BLP* and let $\bar{p}^t = (\bar{x}^t, \bar{u}^t)$. Suppose the current best value of ϕ is k, which may or may not be equal to $\phi(\bar{x}, \bar{y})$.

DEFINITION: Given m positive scalars $\bar{\lambda}_1, \bar{\lambda}_2, \ldots, \bar{\lambda}_m$, then the inequality

$$\sum_{j \epsilon J} p_j / \bar{\lambda}_j \geq 1$$

is a *valid cutting plane* with respect to $\bar{p}^t = (\bar{x}^t, \bar{u}^t)$ if

$$\sum_{j \epsilon J} \bar{p}_j / \bar{\lambda}_j < 1,$$

but

$$\sum_{j \epsilon J} p_j / \bar{\lambda}_j \geq 1$$

for all $p \epsilon P$ such that $\phi(x, y) < k$ for some $y \epsilon Y_0$ where $P = \{(x^t, u^t) : Ex + u = e, \ x \geq 0, \ u \geq 0\}$.

A valid cutting plane thus cuts off the pseudo-global minimum but does not cut off any feasible point of X_0 which along with a $y \epsilon Y_0$ yields an objective function value smaller than k.

The following theorem states how a valid cutting plane can be generated from a pseudo-global minimum.

THEOREM 8: Let $(\bar{x}, \bar{y})$ be a pseudo-global minimum and let the rays ξ^j be as defined in Theorem 6. Let k be the current best value of ϕ. Let $\bar{\lambda}_j$ be defined by

$$\bar{\lambda}_j = \max \{\lambda_j : \bar{x} - \bar{e}^j \lambda_j \epsilon Y^0(k)\} \qquad \text{if } \xi^j \not\subset Y^0(k)$$

$$= \infty \qquad \text{if } \xi^j \subset Y^0(k) \text{ for all } \lambda_j \geq 0.$$

Then the inequality

$$\sum_{j \epsilon J} p_j / \bar{\lambda}_j \geq 1$$

is a valid cutting plane.

PROOF: For notational convenience, let us translate the origin to $(\bar{x}, \bar{y})$. Since $\phi(0, y) \geq \phi(0, 0)$ for all $y \epsilon Y_0$, $0 \epsilon Y^0(k)$. From Theorem 6, X'_0 has precisely m edges incident on 0 and each ξ^j contains one such edge. Let $0 \neq \hat{x} \epsilon X'_0$ be on ray ξ^j. We will show that there exists $\bar{\alpha} > 0$ such that $\phi(\alpha_j \hat{x}, y) \geq 0$ for $0 < \alpha_j \leq \bar{\alpha}$ for all $y \epsilon Y^0$. Suppose, on the contrary

$$\phi(\alpha_j \hat{x}, \hat{y}) = \alpha_j (c^t \hat{x} + \hat{x}^t C \hat{y}) + d^t \hat{y} < 0 \tag{2}$$

for some $\hat{y} \epsilon Y_0$. Since $(0, 0)$ is a pseudo-global minimum, $d^t \hat{y} \geq 0$ by Theorem 5. If $d^t \hat{y} = 0$, by Theorem 5 $(c^t x + x^t C \hat{y}) \geq 0$ for all $x \epsilon X_0$, i.e., $\phi(x, \hat{y}) \geq \phi(0, \hat{y})$ since $d^t \hat{y} = 0$. This implies $x = 0$ solves Min $\phi(x, \hat{y})$, $x \epsilon X_0$. From Theorem 6 then $\phi(x, \hat{y}) = \phi(0, \hat{y})$ is a supporting hyperplane to X'_0 at $x = 0$. That is, $\phi(x, \hat{y}) = c^t x + x^t C \hat{y} \geq 0$ for each $x \epsilon X'_0$ and in particular for $x = \hat{x}$. Thus, $\phi(\alpha_j \hat{x}, \hat{y}) \geq 0$ for all $\alpha_j > 0$, a contradiction. On the other hand, if $d^t \hat{y} > 0$, from the expression for $\phi(\alpha_j \hat{x}, \hat{y})$, we have $\phi(\alpha_j \hat{x}, \hat{y}) \geq 0$ for all

$$0 < \alpha_j \leq \hat{\alpha}_j = \begin{cases} -d^t \hat{y} / (c^t \hat{x} + \hat{x}^t C \hat{y}) & \text{if } (c^t \hat{x} + \hat{x}^t C \hat{y}) < 0 \\ \infty & \text{otherwise} \end{cases}$$

Again this is a contradiction. Now let

$$\bar{\alpha} = \min_j \alpha_j$$

where $\bar{\alpha}$ can be let to be arbitrarily large if all $\alpha_j = \infty$. We have $\phi(\alpha_j \hat{x}, y) \geq 0$ for $0 < \alpha_j \leq \bar{\alpha}$ for all $y \epsilon Y^0$. In other words, $\alpha_j \hat{x} = (\bar{x} - \bar{e}^j \lambda_j) \epsilon Y^0(k)$ with $\lambda_j > 0$ for each $j \epsilon J$. Hence, noting that $Y^0(k)$ is closed, $\bar{\lambda}_j > 0$ exists, i.e., $0 \leq 1/\bar{\lambda}_j < \infty$ and $(\bar{x} - \bar{e}^j \bar{\lambda}_j) \epsilon Y^0(k)$.

To show that the cutting plane generated at $(0, 0)$ is a valid cut, first note that

$$\sum_{j \epsilon J} \bar{p}_j / \bar{\lambda}_j = 0 < 1$$

since the nonbasic variables

$$\bar{p} = \begin{pmatrix} \bar{x} \\ \bar{u} \end{pmatrix}$$

are zero. To complete the proof, consider a feasible

$$p=\begin{pmatrix}\hat{x}\\ \hat{u}\end{pmatrix}$$

such that $\phi(\hat{x}, \hat{y})<$k for some $\hat{y}\epsilon Y_0$. Note that feasibility implies $\hat{p}\geq 0$. We will show that

$$\sum_{j\epsilon J}\hat{p}_j/\bar{\lambda}_j\geq 1.$$

Note that by definition of $\bar{\lambda}_j$,

$$\min_{y\epsilon Y_0}\phi(-\bar{e}^j\bar{\lambda}_j, y)=k \text{ or } \phi(-\bar{e}^j\bar{\lambda}_j, \hat{y})\geq k.$$

Hence, $\hat{x}$ is in one open half space defined by $\phi(x, \hat{y})=k$ and $(-\bar{e}^j\bar{\lambda}_j)$, $j\epsilon J$ are in its complement which is a closed half space. Thus, $\hat{p}$ is feasible to the cut

$$\sum_{j\epsilon J}p_j/\bar{\lambda}_j\geq 1$$

since the hyperplane $\sum p_j/\bar{\lambda}_j=1$ passes through the points $(-\bar{e}^j\bar{\lambda}_j)$. This completes the proof.

In order to define a cutting plane we then need to determine $\bar{\lambda}_j$ specified in Theorem 8. By definition, for $j=1, \ldots, m$

$$\bar{\lambda}_j=\max_{\lambda_j>0}\left\{\min_{y\epsilon Y_0}c^t(\bar{x}-\bar{e}^j\lambda_j)+d^ty+(\bar{x}-\bar{e}^j\lambda_j)^tCy\geq k\right\} \tag{3}$$

This amounts to solving m parametric linear programming problems over Y_0. One way of implementing this will be as follows: let

$$\psi(\lambda_j)=\min_{y\epsilon Y_0}\{c^t(\bar{x}-\bar{e}^j\lambda_j)+d^ty+(\bar{x}-\bar{e}^j\lambda_j)^tCy\}.$$

Since ψ is a concave function of λ_j (see [20]), it is unimodal. We can conduct a search for $\bar{\lambda}_j$ over an interval $(0, L)$ where L is a large enough number, using Bolzano Search [31]. A linear programming problem is solved for a fixed value of λ_j in the interval. Note that we elected to apply the cut only in the x-variables in order to take advantage of ease in solving problems in the y-variables. At each iteration of the search, the "interval of uncertainty" is reduced by 1/2. The search is terminated when a value of $\lambda_j=\hat{\lambda}_j$ is obtained such that $\Psi(\hat{\lambda}_j)-k\leq\beta$ where $\beta>0$ is a prespecified tolerance level. Observe that $\hat{\lambda}_j\leq\bar{\lambda}_j$ and hence the cut using $\hat{\lambda}_j$ is also a valid cut. The search procedures will require the solution of at most $(p+1)$ linear programs where p is the smallest positive integer for which $L/2^p\leq\beta$.

The proposed cutting plane algorithm can be summarized as follows:

1. If the unexplored feasible region $X_0{}^i$ at stage i is empty, terminate. Otherwise, find a pseudo-global minimum. If necessary, update the value of the current best solution.
2. Solve the parametric linear programs to obtain $\bar{\lambda}_j$.
3. Introduce the cut and return to step 1.

Convergence of the algorithm is readily proved. Let $\{x^i, y^i\}$ be the sequence of pseudo-global minima generated and $H(x^i)$ be the cutting plane generated at x^i. At stage i, the algorithm is terminated if $X_0\cap H^+(x^i)=\phi$. Otherwise, the cut $H(x^i)$ is applied and a new pseudo-global point

(x^{i+1}, y^{i+1}) is found where $x^{i+1} \epsilon X_0 \cap H^+(x^i)$ and $x^i \notin H^+(x^i)$. We wish to show that the sequence $\{x^i\}$ has a limit point x^k such that $X_0 \cap H^+(x^k) = \phi$.

Since the points x^i are in a compact set, there exists a limit point x^k such that for a given $\epsilon > 0$ and a positive integer N, $||x^i - x^k|| < \epsilon$ for some $i \geq N$. If $X_0 \cap H^+(x^k) \neq \phi$, all subsequent pseudo-global minima (x^l, y^l) generated will satisfy the condition $x^l \epsilon H^+(x^k)$, $l \geq k+1$. Also, from the proof of Theorem 8 $x^k \epsilon B_\delta(x^k) \cap X^0$ and $x^l \notin xB_\delta(x^k)$ for some $\delta > 0$. Hence, $||x^l - x^k|| > \delta$ for all $l \geq k+1$. This contradicts the statement that x^k is a limit point. Hence, $X^0 \cap H^+(x^k) = \phi$ and the algorithm is terminated.

We can readily show that the cuts generated using generalized polars are uniformly stronger than those generated by Konno [23]. This cut is of the form

$$\sum_{j \epsilon J} d_j p_j \geq \sigma$$

where the d_j are positive constants which are selected and σ is a parameter which is defined to satisfy certain conditions. For each $j \epsilon J$, Konno calculates the maximum value of σ_j such that $(\bar{x} - \bar{e}^j \sigma_j / d_j) \epsilon Y^0(k)$, and then selects $\bar{\sigma}$ such that $\bar{\sigma}/d_j \leq \sigma_j / d_j$ for all $j \epsilon J$. Now $\sigma_j / d_j \leq \bar{\lambda}_j$, hence $\bar{\sigma}/d_j \leq \bar{\lambda}_j$ for all $j \epsilon J$. Hence the cuts generated from $Y^0(k)$ are stronger than those generated by Konno's method.

There is a revealing geometric interpretation to this difference. Konno predetermines the coefficients d_j of the cutting plane so as to simplify computational work. But this has the effect fo fixing the slope of the hyperplane. The hyperplane is now translated parallel to itself till such time as one point on it touches the boundary of $Y^0(k)$. The polar cut allows the additional flexibility of altering the slope of the cutting plane so as to cut off more of the feasible region.

V. A FINITELY CONVERGENT ϵ-OPTIMAL ALGORITHM

The above algorithm can be converted to a finitely convergent ϵ-optimal solution procedure by using $Y^0(k\text{-}\epsilon)$ instead of $Y^0(k)$ in determining $\bar{\lambda}_j$. In this case, a cut can be initiated at a local star minimum rather than a pseudo-global minimum. Note that from Theorem 1,

$$\phi(0, y) \geq \phi(0, 0) \geq k > k - \epsilon,$$

for each $y \epsilon Y_0$. Hence, 0ϵ int $Y^0(k - \epsilon)$. This implies that there exists a $\bar{\alpha}$ such that $\phi(\alpha_j \hat{x}, y) \geq k$ for $0 < \alpha_j \leq \bar{\alpha}$ and for all $y \epsilon Y_0$ where $\hat{x}$ is as defined in the proof of Theorem 8. The remaining parts of Theorem 8 hold. Hence, a valid cut can be generated from a local star minimum.

We will now show that the algorithm is finite. At stage i, suppose the local star minimum is $(\bar{x}, \bar{y})$ and let k_i be the current best value of ϕ attained at $(\hat{x}, \hat{y})$, which may or may not be the same as $\phi(\bar{x}, \bar{y})$. Then $(\hat{x}, \hat{y})$ is the ϵ-optimal solution over all $y \epsilon Y_0$ and over X_0 already explored, i.e., for all x in X_0, but not in X_0^{i+1}. Note that k_i is decreased at each stage at least by a fixed $\epsilon > 0$. Hence, denoting by $(\bar{\bar{x}}, \bar{\bar{y}})$ the global optimum of *BLP*, if we show that $\bar{\bar{x}}$ is cut off by a cut obtained from $Y^0(\bar{\bar{k}})$ for some $\bar{\bar{k}}$, then the algorithm is finite.

Now consider *BLP* with $(\bar{x}, \bar{y})$ as the local star minimum. If $(\bar{\bar{x}}, \bar{\bar{y}}) \epsilon X_0^{i+1}$ is the global minimum, then $\bar{\bar{x}}$ is in the cone with vertex $\bar{x}$ and ξ^j, $j \epsilon J$, as the generators. Then $\bar{\bar{x}}$ can be expressed as a convex combination of points $\bar{x}^j$, $j \epsilon J$, on these generators.

Let

$$\alpha_j = \min_{y \in Y_0} \phi(\bar{x}^j, y)$$

and

$$\bar{\bar{k}} = \min \alpha_j.$$

Then $Y^0(\bar{\bar{k}})$ will cut the rays ξ^j at points $(0, \ldots, \bar{\lambda}_j, \ldots, 0) \geq \bar{x}^j$. Hence,

$$\sum_{j \in J} p_j/\bar{\lambda}_j \geq 1$$

obtained from $Y^0(\bar{\bar{k}})$ will cut off $\bar{\bar{x}}$. This shows that the ϵ-optimal algorithm is finite.

VI. GENERATION OF DEEPER CUTS

We are grateful to the referee for bringing to our attention some recent work of Burdet [8, 9] dealing with convex gauge functions which permits generation of deeper cuts than that given in Section IV. This approach allows $\bar{\lambda}_j$ to be negative and generates uniformly dominating cuts. Owen [25] seems to be the first to suggest this aproach, and related work on negative edge extensions has been investigated in [4, 16, 18, 22]. Using these results the cut

$$\sum_{j \in J} p_j/\bar{\lambda}_j \geq 1$$

can be strengthened by using the following definition for $\bar{\lambda}_j$.

$$\lambda_j = \begin{cases} \max_{\lambda_j > 0} \{\lambda_j < \infty : \phi(\bar{x} - \bar{e}^j \lambda_j, y) \geq k \text{ for all } y \in Y^0\} \text{ if } \xi^j \not\subset Y^0(k) \\ -[\max_{\lambda_j > 0} \{\lambda_j \leq \infty : \phi(\bar{x} + \bar{e}^j \lambda_j, y) \geq k \text{ for some } y \in Y^0\}] \text{ if } \xi^j \subset Y^0(k) \end{cases}$$

If $\xi^j \not\subset Y^0(k)$, $0 < \bar{\lambda}_j < \infty$ can be computed from equation (3) given earlier by solving parametric transportation problems. If $\xi^j \subset Y^0(k)$ as pointed out by the referee, the value of $\bar{\lambda}_j < 0$ can be computed by solving parametric problems similar to that used for (3). More specifically, in this case we need to solve the parametric problem:

$$\bar{\lambda}_j = -[\max_{\lambda_j < 0} \{\lambda_j : \max_{y \in Y_0} \phi(\bar{x} + \bar{e}^j \lambda_j, y) \geq k\}].$$

REFERENCES

[1] Balas, E., "Intersection Cuts—A New Type of Cutting Planes for Integer Programming," Operations Research, *19*, 19–39 (1971).

[2] Balas, E., "Integer Programming and Convex Analysis: Intersection Cuts from Outer Polars," Mathematical Programming, *2*, 330–382 (1972).

[3] Balas, E., "Nonconvex Quadratic Programming Via Generalized Polars," Management Science Research Report No. 278, GSIA, Carnegie-Mellon University (1972).

[4] Balas, E., "Disjunctive Programming: Properties of the Convex Hull of Feasible Points," Management Science Research Report 348, GSIA, Carnegie-Mellon University (1974).

[5] Balas, E., and C. A. Burdet, "Maximizing a Convex Quadratic Function Subject to Linear Constraints," Management Science Research Report No. 299, GSIA, Carnegie-Mellon University (1973).

[6] Burdet, C. A., "Polaroids: A New Tool in Nonconvex and in Integer Programming," Naval Research Logistics Quarterly, *20*, 13–22 (1973).

[7] Burdet, C. A., "On Polaroid Intersections" in Mathematical Programming in Theory and Practice, P. Hammer and G. Zoutendijk, eds. (North Holland, 1974).

[8] Burdet, C. A., "Elements of a Theory in Nonconvex Programming," to appear in Naval Research Logistics Quarterly.

[9] Burdet, C. A., "Convex and Polaroid Extensions," to appear in Naval Research Logistics Quarterly.

[10] Cabot, V. A., and R. L. Francis, "Solving Certain Nonconvex Quadratic Minimization Problems by Ranking the Extreme Points," Operations Research, *18*, 82–86 (1970).

[11] Candler, W. and R. J. Townsley, "The Maximization of a Quadratic Function of Variables Subject to Linear Inequalities," Management Science, *10*, 515–523 (1964).

[12] Cottle, R. W., and W. C. Mylander, "Ritter's Cutting Plane Method for Nonconvex Quadratic Programming," in *Integer and Nonlinear Programming*, J. Aladie, ed., (North Holland, 1970).

[13] Gallo, G., and A. Ülkücü, "Bilinear Programming: An Exact Algorithm," Report ORC 73–26, Operations Research Center, University of California, Berkeley (1973).

[14] Geoffrion, A. M., "Elements of Large-Scale Mathematical Programming," Management Science, *16*, 652–691 (1970).

[15] Glover, F., "Convexity Cuts and Cut Search," Operations Research, *21*, 123–134 (1973).

[16] Glover, F., "Polyhedral Convexity Cuts and Negative Edge Extensions," Zeitschrift für Operations Research, *18*, 181–186 (1974).

[17] Glover, F., "Convexity Cuts for Multiple Choice Problems," Discrete Mathematics, *6*, 221–234 (1973).

[18] Glover, F., "Polyhedral Annexation in Mixed Integer and Combinational Programming," Mathematical Programming, *8*, 161–188 (1975).

[19] Glover, F., and D. Klingman, "Concave Programming Applied to a Special Class of 0–1 Integer Programs," Operations Research, *21*, 135–140 (1973).

[20] Hillier, F. S., and G. J. Lieberman, *Introduction to Operations Research* (Holden-Day, 1974).

[21] Hu, T. C., "Minimizing a Concave Function in a Convex Polytope," Mathematics Research Center Report No. 1011, U.S. Army, Madison, Wisconsin (1969).

[22] Jeroslow, R. G., "The Principles of Cutting-Plane Theory: Part I," with an addendum GSIA, Carnegie-Mellon University (1974).

[23] Konno, H., "Bilinear Programming," Parts I and II. Technical Report No. 71–9 and 71–10, Operations Research House, Stanford University (1971).

[24] Mueller, R. K., "A Method for Solving the Indefinite Quadratic Programming Problem," Management Science, *16*, 333–339 (1970).

[25] Owen, G., "Cutting Planes for Programs with Disjunctive Constraints," Journal of Optimization Theory and Applications, *11*, 49–55, 1973.

[26] Ritter, K., "A Method for Solving Maximum-Problems with a Nonconcave Quadratic Objective Function," Z. Wahrscheinlichkeitstheorie verw., *4*, 340–351 (1966).

[27] Tuy, H., Concave Programming Under Linear Constraints (Russian) *Doklady Akademii Nauk SSSR* (1964). English translation in Soviet Mathematics, *5*, 1437–1440 (1964).

[28] Vaish, H., "Nonconvex Programming with Applications to Production and Location Problems," Unpublished Ph.D. Dissertation, Georgia Institute of Technology (1974).

[29] Vaish, H., and C. M. Shetty, "The Bilinear Programming Problem," Naval Research Logistics Quarterly, *23*, No. 2 (1976).

[30] Young, R. D., "Hypercyclindrically Deduced Cuts in Zero-One Integer Programming," Operations Research, *19*, 1393–1405 (1971).

[31] Zangwill, W. I., *Nonlinear Programming—A Unified Approach*, (Prentice-Hall, 1969).

THE EFFECT OF CORRELATED EXPONENTIAL SERVICE TIMES ON SINGLE SERVER TANDEM QUEUES

C. R. Mitchell
U.S. Air Force Academy

A. S. Paulson
Rensselaer Polytechnic Institute

C. A. Beswick
University of South Carolina

ABSTRACT

An investigation via simulation of system performance of two stage queues in series (single server, first-come, first-served) under the assumption of correlated exponential service times indicates that the system's behavior is quite sensitive to departures from the traditional assumption of mutually independent service times, especially at higher utilizations. That service times at the various stages of a tandem queueing system for a given customer should be correlated is intuitively appealing and apparently not at all atypical. Since tandem queues occur frequently, e.g. production lines and the logistics therewith associated, it is incumbent on both the practitioner and the theoretician that they be aware of the marked effects that may be induced by correlated service times. For the case of infinite interstage storage, system performance is improved by positive correlation and impaired by negative correlation. This change in system performance is reversed however for zero interstage storage and depends on the value of the utilization rate for the case where interstage storage equals unity. The effect due to correlation is shown to be statistically significant using spectral analytic techniques. For correlation equal unity and infinite interstage storage, results are provided for two through twenty-five stages in series to suggest how adding stages affects system performance for $\rho > 0$. In this extreme case of correlation, adding stages has an effect on system performance which depends markedly on the utilization rate. Recursive formulae for the waiting time per customer for the cases of zero, one, and infinite interstage storage are derived.

1. INTRODUCTION

First, we describe two physical settings where dependent service times can be expected to arise. In a paper mill, large rolls of paper typically pass through an inspection or winding operation prior to being cut into smaller rolls. A poor quality roll takes a relatively longer time in the inspection process because defective sections must be removed and splices made. When this same roll reaches the final cutting stage it must be processed more slowly to avoid breaking the splices and to repair them when they do break. Hence process times at the two stages tend to be correlated; indeed, it is conceptually possible that they be highly correlated. The process times at the two stages on any

two different rolls would generally be independently distributed. In the current context considerable interest would be centered on the effect, if any, produced by nonindependence of process times at different stages.

Jackson [9], in discussing queueing systems with phase type service, pointed out that a typical sequence of events in the overhaul of an aircraft engine consists of stripping, detailed examination, repairs, assembly, and testing. Generally, an engine with a large number of maintenance requirements can be expected to spend more time in each of the latter four phases and so the possible effect of correlated service times on throughput time would be of interest. It is not difficult to envision a host of other situations involving queues in series in which the service times at the various stages for a given customer are correlated.

A large proportion of the literature concerning tandem queues has centered on Poisson arrival processes, exponential service times, and steady state solutions. The assumption of independence of service times is intricately interwoven into the fabric of the traditional birthdeath equation approach to finding a transient and steady state solution to the tandem queueing phenomenon. We shall remain within this same framework with the exception that we shall drop the heretofore universal (but tacit!) assumption of mutual independence of all exponential service times. An obvious approach is to use a multivariate exponential distribution with non-zero correlations in place of the usual independent exponential service times. In our situation it is not clear that the birthdeath equation approach can be modified to incorporate dependent service times. Moreover, any such formulation would very likely be analytically intractable. The problem is, however, amenable to a simulation approach and it is in this way that we assess the effect of departures from independence of service times on steady state system performance.

2. THE SERIES QUEUEING SYSTEM AND RECURSIVE FORMULAE

Consider the series queueing process depicted in Figure 1. Customers from an infinite population arrive at a two stage system according to a Poisson process with mean rate λ which we shall, without loss of generality, take to be unity. An unlimited queue is always allowed before the first stage, but before the second stage the queue length may be either restricted or unlimited. A single server is allowed at each stage; the service discipline is first-come, first-served.

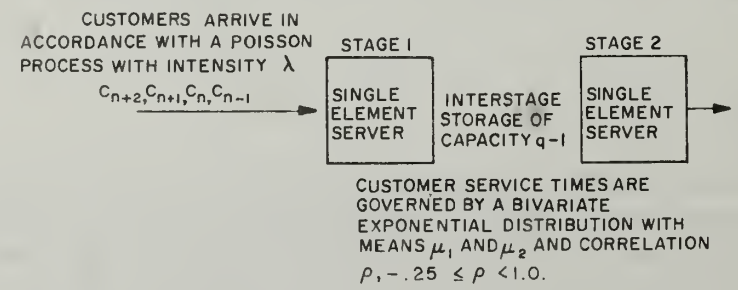

CUSTOMERS ARRIVE IN ACCORDANCE WITH A POISSON PROCESS WITH INTENSITY λ
$C_{n+2}, C_{n+1}, C_n, C_{n-1}$
STAGE 1
SINGLE ELEMENT SERVER
INTERSTAGE STORAGE OF CAPACITY q-1
STAGE 2
SINGLE ELEMENT SERVER
CUSTOMER SERVICE TIMES ARE GOVERNED BY A BIVARIATE EXPONENTIAL DISTRIBUTION WITH MEANS μ_1 AND μ_2 AND CORRELATION $\rho, -.25 \leq \rho < 1.0$.

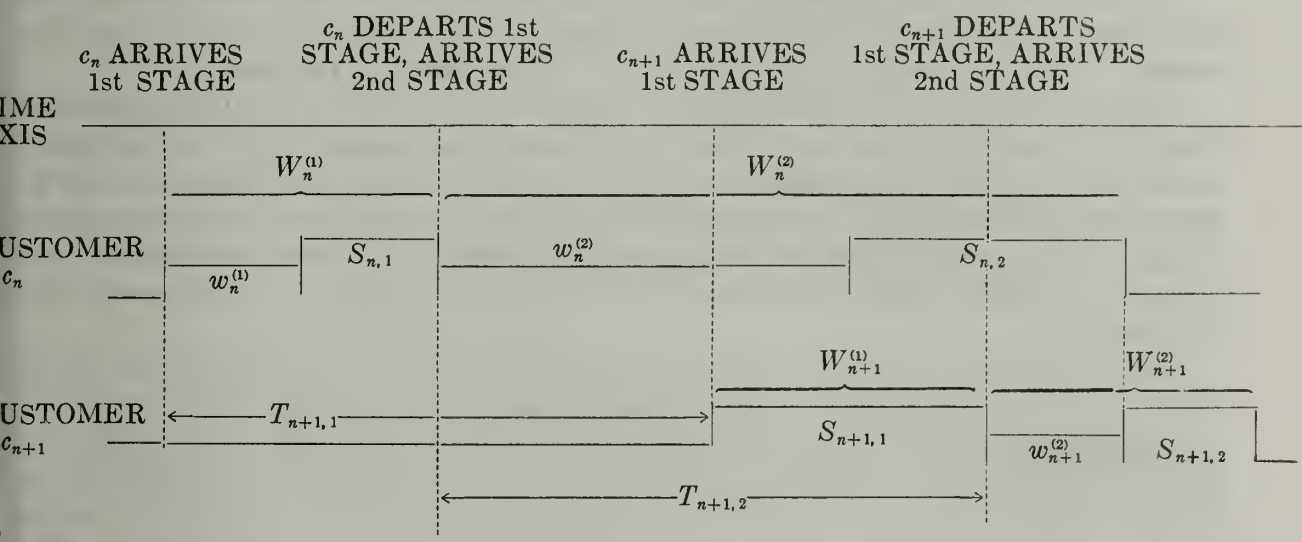
FIGURE 1.—Two stage series queue with dependent service times.

The system performance measure is taken to be mean waiting time per customer and in this section we develop a set of formulae to recursively compute the waiting time per customer. We use the recursive formulae for the unlimited interstage storage case in order to demonstrate precisely how the queueing system consisting of two stages in series with dependent service times is

related to a single server system with interdependent arrival and service processes as discussed by Bhat [2]. An interpretation of Conolly [4] for a special type of this latter interdependence is shown to be helpful in suggesting why mean waiting time is affected by correlated service times.

Denote by $(T_{n,1}, T_{n,2})$ the times between arrival epochs of customers c_{n-1} and c_n at the first and second stages and let c_n experience the service times $(S_{n,1}, S_{n,2})$ at each stage, $n=1, 2, \ldots$. The sequences of interarrival times $(T_{n,1}, T_{n,2})$ and the $(S_{n,1}, S_{n,2})$ for different customers are both assumed to be mutually independent.

We take $(w_n^{(1)}, w_n^{(2)})$ to be the waiting times, excluding service, and $(W_n^{(1)}, W_n^{(2)})$ to be the total waiting times, of customer c_n, at the respective stages, $n=1, 2, \ldots$. We illustrate these definitions with an arbitrary combination of arrival and service times in the following diagram. The illustration is for two queues in series with unlimited interstage storage; diagrams like this are useful in developing the recursive formulae for the different cases to be presented.



CASE A: Two stage queues in series, unlimited interstage storage.

Customer c_{n+1}'s total waiting time at the first stage and interarrival time at the second stage are given by

$$W_{n+1}^{(1)} = \begin{cases} S_{n+1,1}, & \text{if } T_{n+1,1} \geq W_n^{(1)} \\ W_n^{(1)} - T_{n+1,1} + S_{n+1,1}, & \text{if } T_{n+1,1} < W_n^{(1)} \end{cases} \tag{1}$$

and

$$T_{n+1,2} = \begin{cases} T_{n+1,1} - W_n^{(1)} + S_{n+1,1}, & \text{if } T_{n+1,1} \geq W_n^{(1)} \\ S_{n+1,1}, & \text{if } T_{n+1,1} < W_n^{(1)} \end{cases}. \tag{2}$$

The condition in (1) and (2) that $T_{n+1,1} \geq W_n^{(1)}$ simply means that c_{n+1} arrives at server one after c_n has departed, and likewise $T_{n+1,1} < W_n^{(1)}$ means c_{n+1} arrives before c_n leaves.

Similar to (1), c_{n+1}'s waiting time at the second stage is

$$W_{n+1}^{(2)} = \begin{cases} S_{n+1,2}, & \text{if } T_{n+1,2} \geq W_n^{(2)} \\ W_n^{(2)} - T_{n+1,2} + S_{n+1,2}, & \text{if } T_{n+1,2} < W_n^{(2)} \end{cases}. \tag{3}$$

The above diagram illustrates (1), (2), and (3) for $T_{n+1,1} \geq W_n^{(1)}$ and $T_{n+1,2} < W_n^{(2)}$. Similar diagrams result for the remaining conditions.

In an obvious way, we can use these relationships to build up a set of recursive formulae for any number of stages in series where the interstage storage between stages is unlimited. (See Appendix.)

Since each customer must proceed through both stages, the output of the first stage becomes the input of the second stage and therefore we have, in steady state, that the time interval between arrivals at the second stage satisfies a Poisson process with the same interarrival intensity parameter λ as the input distribution [3]. Unlike the first stage however, c_n's service time at the second stage is correlated with the interarrival time there. In the above diagram, this corresponds to a correlation between $S_{n+1,2}$ and $T_{n+1,2}$. This result is apparent from (2) since $S_{n+1,1}$ and $S_{n+1,2}$ are dependent by assumption.

If $S_{n+1,1}$ and $S_{n+1,2}$ are independent as is usually assumed for two stage server systems, then each stage, in steady state, can be analyzed independently, and since $T_{n+1,2}$ and $S_{n+1,2}$ are independent, as are $T_{n+1,1}$ and $S_{n+1,1}$, the regular M/M/1 results obtain for each stage.

Bhat [2] describes five different classes of single server first-come first-served systems with Poisson input and exponential service times which result from relaxing some of the assumptions of independence which are typically made. These classes represent more realistic operating systems than those with assumptions of independence; Bhat further points out that more work needs to be done on these problems than the limited amount reported at that time (1969). One of these classes is for systems with interdependent arrival and service processes as is the case here for $T_{n+1,2}$ and $S_{n+1,2}$.

Conolly [4] and Conolly and Hadidi [5] have studied a dependent structure somewhat similar to this wherein the ratio of service time to interarrival time is constant for all n; they give transient as well as steady state results for the system. Conolly showed numerically that this pattern of server behavior results in a drastic reduction in the mean and variance of the waiting time as compared with a conventional M/M/1 queue. It was noted by Conolly that this kind of server behavior is to be expected from a well regulated service facility where the server adjusts the service time of a customer according to that customer's interarrival time, which the server observes without error. In this way, a long interval gives rise to a long service time, and short intervals corresponding to a succession of rapid arrivals are followed by correspondingly short service times. This regulated behavior therefore prevents a long queue from forming and cuts down on the mean and variance of the waiting time in the system.

Returning to the two stages in series problem under study we see that this system, via equation (2), can be viewed as a type of self-regulated system since $S_{n+1,\,2}$ and $T_{n+1,\,2}$ are related, although not in the deterministic way assumed by Conolly. It will be demonstrated later that our type of stochastic dependence between $S_{n+1,\,2}$ and $T_{n+1,\,2}$ gives rise to results which are consistent with Conolly's. This artificial way of viewing the system as a self-regulating device is employed solely to make the effects seem more reasonable and in no way influences the results.

CASE B: Two stage queues in series, no interstage storage.

For this case, c_n's total waiting time at the second stage, $W_n^{(2)}$, is always equal to the $S_{n,\,2}$ so the only quantity of interest here is $W_n^{(1)}$. Since there is restricted (zero) interstage storage, the phenomenon of blocking occurs and so the waiting time computation is a bit more complicated than in Case A. (In effect, the first server's utilization is diminished [17].)

The total wating time for c_{n+1} at the first server is given by one of four relationships depending upon whether or not c_{n+1} arrives at stage one before or after c_n leaves.

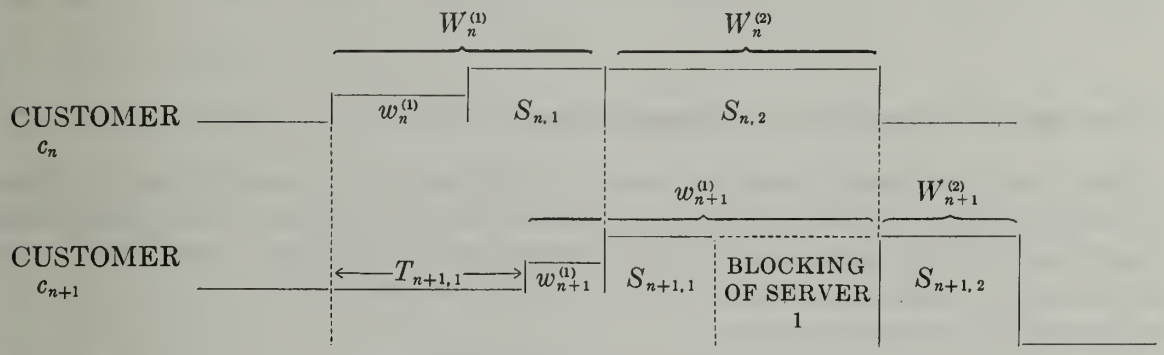
For $T_{n+1,\,1} < W_n^{(1)}$,

$$W_{n+1}^{(1)} = \begin{cases} W_n^{(1)} - T_{n+1,1} + S_{n+1,1}, & \text{if } S_{n+1,1} \geq S_{n,2} \\ W_n^{(1)} - T_{n+1,1} + S_{n,2}, & \text{if } S_{n+1,1} < S_{n,2} \end{cases} \tag{4}$$

and for $T_{n+1,\,1} \geq W_n^{(1)}$,

$$W_{n+1}^{(1)} = \begin{cases} S_{n+1,1}, & \text{if } S_{n+1,1} \geq W_n^{(1)} - T_{n+1,1} + S_{n,2} \\ W_n^{(1)} - T_{n+1,1} + S_{n,2}, & \text{if } S_{n+1,1} < W_n^{(1)} - T_{n+1,1} + S_{n,2} \end{cases} \tag{5}$$

The following diagram illustrates (4) for $S_{n+1,\,1} < S_{n,\,2}$.



Similarly, the other conditions can be verified.

CASE C: Two stage queues in series, interstage storage capacity equals one.

As in the previous case blocking can occur at the first stage but here a customer's total waiting time at stage two can exceed the service time since interstage storage is permitted on a restricted basis.

If $T_{n+1,1} < W_n^{(1)}$,

$$W_{n+1}^{(1)} = \begin{cases} W_n^{(1)} - T_{n+1,1} + S_{n+1,1}, & \text{if } S_{n+1,1} \geq W_n^{(2)} - S_{n,2} \\ W_n^{(1)} - T_{n+1,1} + W_n^{(2)} - S_{n,2}, & \text{if } S_{n+1,1} < W_n^{(2)} - S_{n,2} \end{cases} \tag{6}$$

and c_{n+1}'s interarrival time at server two is

$$T_{n+1,2} = \begin{cases} S_{n+1,1}, & \text{if } S_{n+1,1} \geq W_n^{(2)} - S_{n,2} \\ W_n^{(2)} - S_{n,2}, & \text{if } S_{n+1,1} < W_n^{(2)} - S_{n,2} \end{cases} \tag{7}$$

If $T_{n+1,1} \geq W_n^{(1)}$,

$$W_{n+1}^{(1)} = \begin{cases} S_{n+1,1}, & \text{if } S_{n+1,1} \geq W_n^{(1)} - T_{n+1,1} + W_n^{(2)} - S_{n,2} \\ W_n^{(1)} - T_{n+1,1} + W_n^{(2)} - S_{n+2}, & \text{if } S_{n+1,1} < W_n^{(1)} - T_{n+1,1} + W_n^{(2)} - S_{n,2} \end{cases} \tag{8}$$

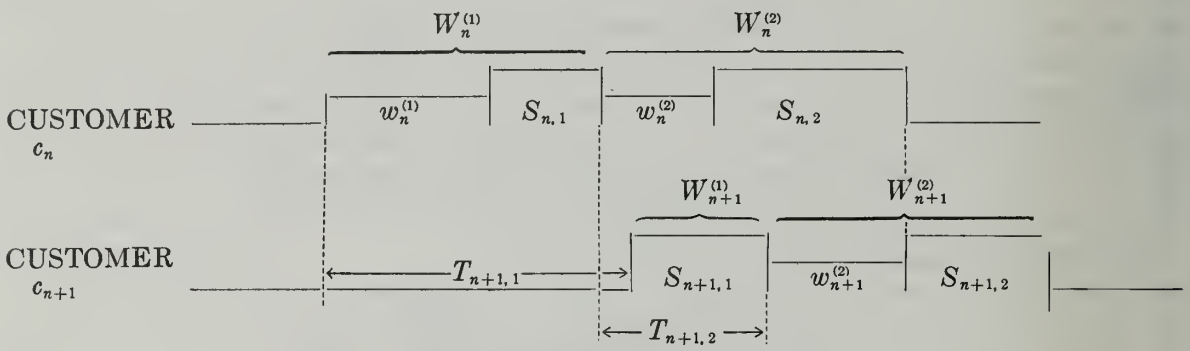
and

$$T_{n+1,2} = \begin{cases} T_{n+1,1} - W_n^{(1)} + S_{n+1,1}, & \text{if } S_{n+1,1} \geq W_n^{(1)} - T_{n+1,1} + W_n^{(2)} - S_{n,2} \\ W_n^{(2)} - S_{n,2}, & \text{if } S_{n+1,1} < W_n^{(1)} - T_{n+1,1} + W_n^{(2)} - S_{n,2} \end{cases} \tag{9}$$

Next the total waiting time of c_{n+1} at stage two, $W_{n+1}^{(2)}$, is computed by using (3) in Case A with $T_{n+1,2}$ as defined in (7) or (9).

The following diagram is descriptive of (8) and (9) where $T_{n+1,1} \geq W_{n+1}^{(1)}$ and

$$S_{n+1,1} \geq W_n^{(1)} - T_{n+1,1} + W_n^{(2)} - S_{n,2}.$$



3. A BIVARIATE EXPONENTIAL DISTRIBUTION

There are a number of bivariate exponential distributions which could be used to describe the dependence assumed between S_1 and S_2 (we drop the subscripts n for now). We choose to use a special case of the bivariate gamma distribution discussed by Wicksell [18] and Kibble [12] and more generally by Krishnamoorthy and Parthasarty [14] and Paulson [16].

The functional form can be written as

$$f(s_1, s_2) = \frac{a}{\theta_1\theta_2} e^{-s_1/\theta_1 - s_2/\theta_2} I_0\left(\left(2\frac{ds_1s_2}{\theta_1\theta_2}\right)^{\frac{1}{2}}\right) \tag{10}$$

where $s_1 \geq 0$, $s_2 \geq 0$, and

$$I_0(z) = \sum_{k=0}^{\infty} \frac{(z/2)^{2k}}{k!k!}$$

is the modified Bessell function of the first kind and order zero. Here $a > 0$, $d \geq 0$, and $a + d = 1$. (See Downton [7])

The actual simulation of variates

$$\underline{S} = \begin{pmatrix} S_1 \\ S_2 \end{pmatrix}$$

from the distribution (10) and its generalization due to Paulson [16] is affected via

$$\underline{S}_\infty = X_1 + V_1X_2 + V_1V_2X_3 + \cdots; \tag{11}$$

here X_j is a 2— vector of independent exponential variates with mean vector

$$\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}$$

and the V_j are random 2×2 matrices which take on values in the set

$$\left\{\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right\}$$

with probabilities a, b, c, d respectively. All the X_j's and V_j's are mutually independent. Note that eventually the product ΠV_j will result in a matrix of zeroes and so with probability one S_∞ is represented by a finite sum [11, 13].

The bivariate random variable $\underline{S}_\infty$ in (11) has mean vector

$$\underline{\mu}=\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}=\begin{pmatrix}\theta_1/(a+c)\\ \theta_2/(a+b)\end{pmatrix}, \tag{12}$$

and covariance matrix

$$\sum=\begin{pmatrix}(\mu_1)^2 & \dfrac{ad-bc}{1-d}\,\theta_1\theta_2\\ \dfrac{ad-bc}{1-d}\,\theta_1\theta_2 & (\mu_2)^2\end{pmatrix}, \tag{13}$$

and for the correlation ρ, $-2.5\leq\rho<1$.

4. SIMULATION RESULTS AND INTERPRETATION

Simulated results are presented in this section for three cases of interstage storage capacity: (A) infinite ($q=\infty$), (B) zero ($q=1$), and (C) one ($q=2$). For the infinite interstage storage case results will be given for two stages in series for various values of correlation and for two through twenty-five stages in series for correlation equal unity. The latter depicts how adding stages might affect system performance given correlation $\rho>0$; more precisely, it provides an envelope within which system performance will vary since for a fixed number of stages and utilization correlation unity provides an extremum and correlation zero provides another. In each of the three cases we allow infinite storage before stage one. In the ordinary case in which the correlation between paired service times is zero, a few steady state results are available for comparison purposes.

We have taken the mean arrival rate to be unity and so the steady state utilization, ν, at stage i is simply the mean service time μ_i. It will suffice for our purposes to take $\mu_1=\mu_2=\mu$ since similar steady state behavior will obtain for $\mu_1\neq\mu_2$. Furthermore, there do not seem to be many results available for puposes of comparison for Cases B and C when $\mu_1\neq\mu_2$. For $\lambda=1$, our system performance measure of mean waiting time (queueing plus service), is equivalent to the expected number in the system.

CASE A: k stage queues in series, infinite interstage storage. (Graphs are labeled k Q for k-queues).

Steady state results for k stages in series with no correlation between pairs of service times are available [17] and we have that the expected number of customers at each stage is $\nu/(1-\nu)$ and $k\nu/(1-\nu)$ in the system.

Figure 2 provides for two stages in series the mean waiting time at the second stage, $\overline{W}_n^{(2)}$, for $\nu=0.75$ and $\rho=-0.25$, 0, 0.50, and 1.0. In this case the mean waiting time at the first stage is independent of ρ since no blocking occurs, and hence it suffices to examine the mean waiting time at the second stage to determine the effects of correlated service times. In some of the simulation results to follow we replicate, many times, runs of much shorter length; here we choose to illustrate the mean waiting time as a function of n with one very long run. Long runs, such as this one, may be considered as being composed of replications of smaller runs where the starting condition of a new

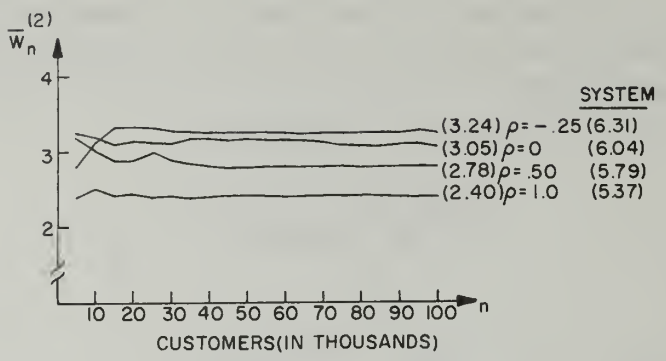


FIGURE 2.—Mean waiting time at second stage, 2Q, $q=\infty$, $\nu=0.75$.

replication is the ending value of the previous replication [6]. From the figure it is clear that each graph tends to stabilize for increasing n in accordance with the law of large numbers.

The numbers adjacent to the values of ρ in Figure 2 are the mean waiting times for the second stage and mean waiting times in the system after 100,000 service completions. We point out that for $\rho=0$ the mean waiting times at stages one and two are 2.99 and 3.05 respectively, and these are in close agreement with the expectation of 3.0 for this utilization. For $\rho\neq 0$, we see here a bonus attached to positive correlation in service times since system performance improves with increasing correlation. On the other hand, system performance deteriorates with negative correlation.

Each figure like this one has a starting condition based on the mean waiting time from a pilot run and then we omit the waiting times of the first 1000 customers in the actual computations shown.

Figure 3 gives system performance for different values of ν and for $\rho=0$, 1. These graphs are intended to show that there is no discernable effect due to correlation $\rho>0$ at utilization $\nu=0.6$ but as ν increases from 0.6 to 0.9 a definite trend appears.

Figure 4 shows the ratio of mean time in the system for various values of ρ to the mean waiting time in the system at $\rho=0$. These kinds of graphs are based on an average of 100 replications of 1400 service completions (after an initialization of 400 service completions were discarded).

The solid lines in Figures 4, 7, and 9 depict a smoothed fit to the actual data. Sampling variation, of course, precludes the possibility of obtaining such a smooth fit without extremely long runs or extensive replications, but each curve was spot-checked to ascertain whether or not the fit was spurious. In no case was any substantial deviation recorded.

Now we show how these effects are consistent with Conolly's [4] results for the case of $\nu=0.9$ and correlation of $\rho=1.0$ between the service times in the two stages. Conolly showed for his single server queueing system where the ratio of service time to the interarrival time was constant for all n, that for a utilization of 0.9 (the ratio) the mean waiting time (queueing plus service) was 2.71. For service time independent of interarrival time the steady state expectation, for this utilization, is 9.0. The interarrival time and service time in Conolly's system are perfectly correlated, whereas in our system the two service times are perfectly correlated. It is clear from equation (2) that the correlation between the interarrival time at the second stage and the service time there is less than one, and so the improvement in system performance for our system should be less than Conolly's (an elementary derivation shows the correlation to be $\nu\rho$ or 0.9 in this case). We see from Figure 3

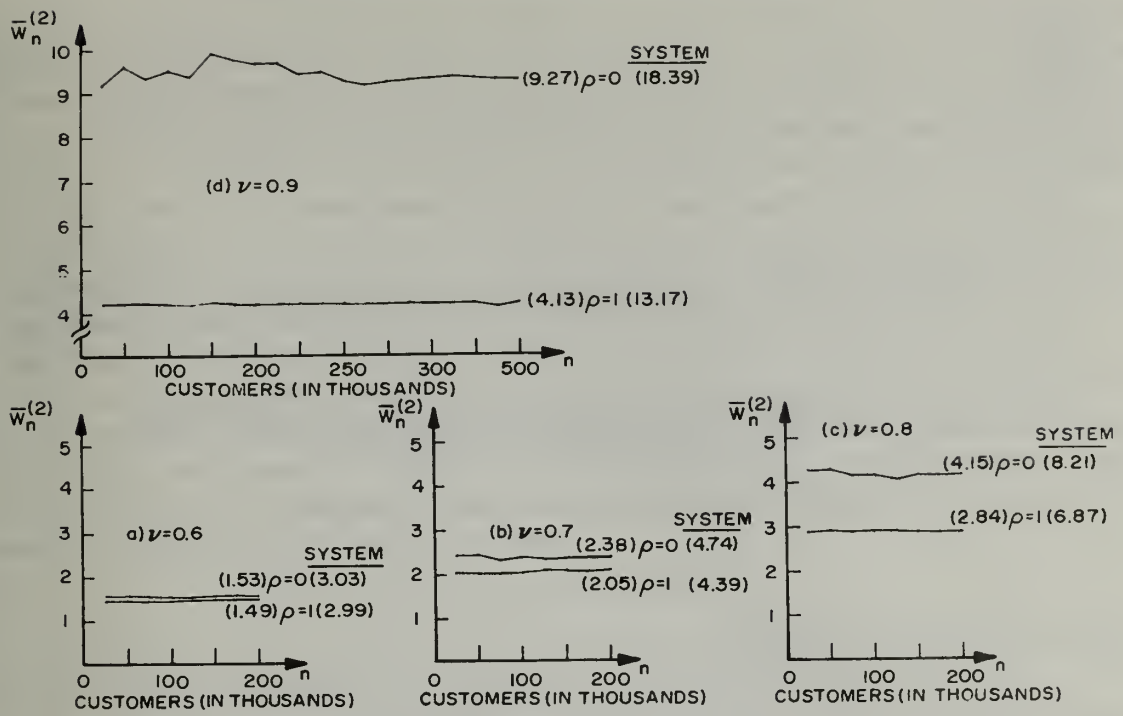


FIGURE 3.—Mean waiting time at second stage, 2Q, $q = \infty$.

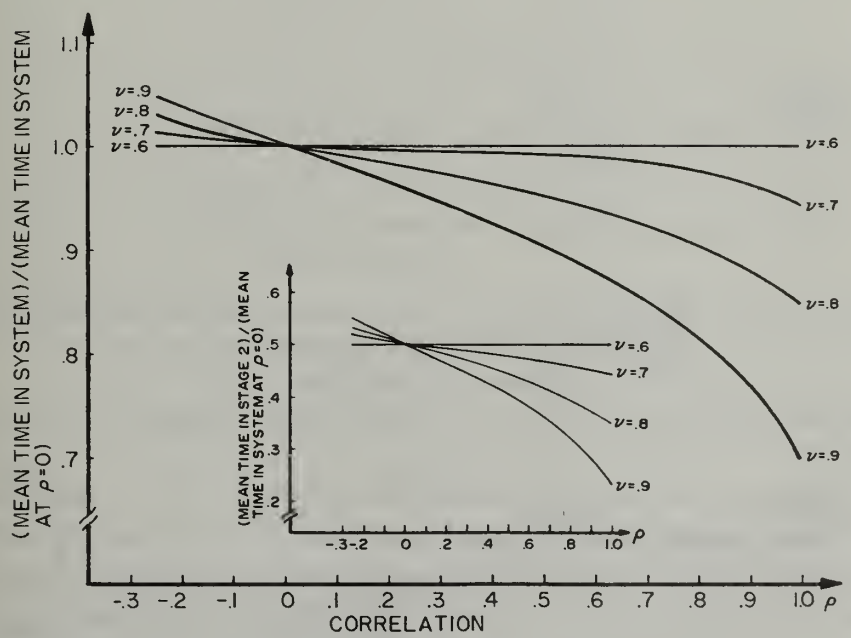


FIGURE 4.—Ratio of mean waiting times, 2Q, $q = \infty$.

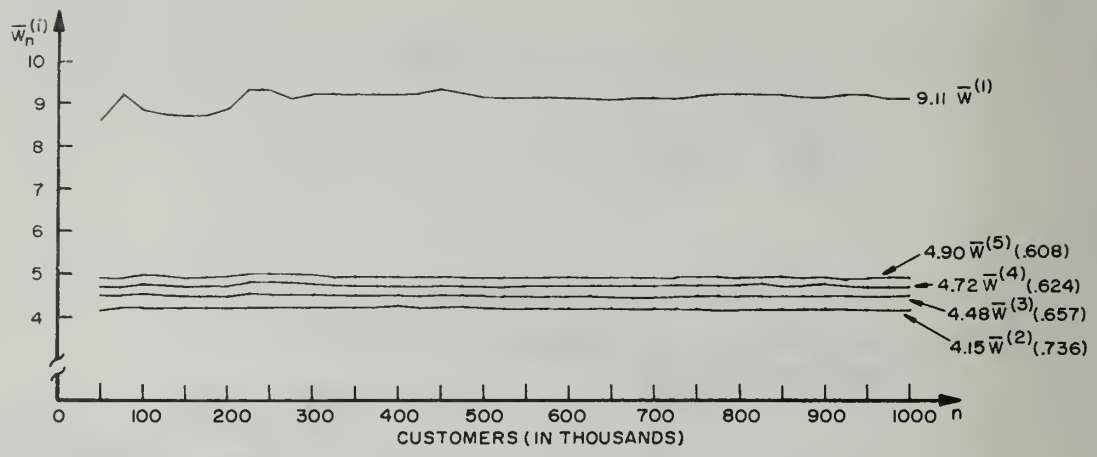


FIGURE 5a.—Mean waiting time, 5Q, $q=\infty$, $\rho=1.0$, $\nu=0.9$.

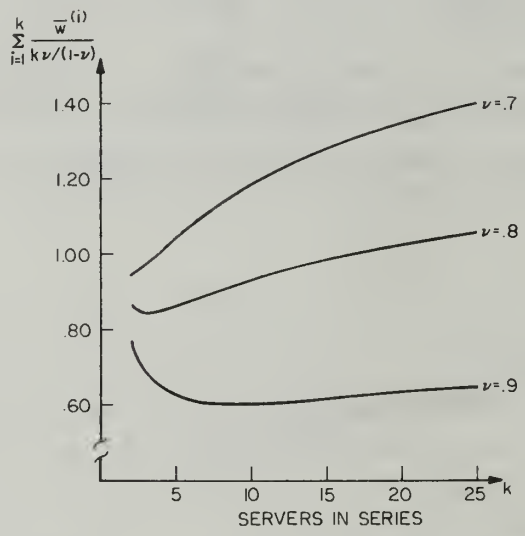


FIGURE 5b.—Ratio of mean waiting time with $\rho=1$ to expected mean waiting time with $\rho=0$, 2Q–25Q, $q=\infty$ (based on 100,000 service completions).

that the mean waiting time at the second stage after 100,000 customers is 4.13, and indeed the improvement is less.

In Figure 5a we show the mean waiting time as a function of n for five stages in series where the service times are equal at each stage. The graphs are labeled $\overline{W}^{(k)}$ corresponding to the mean waiting time at stage k, $k=1, 2, \ldots, 5$. We see that the mean waiting time $\overline{W}^{(2)}$, for the second stage, is consistent with the results in Figure 3. The results for $\overline{W}^{(3)}$, $\overline{W}^{(4)}$, and $\overline{W}^{(5)}$ suggests that further improvements in system performance occur over the $\rho=0$ case but the effect seems to approach a limit. The number in parenthesis to the right of $\overline{W}^{(k)}$ is the ratio

$$\sum_{i=1}^{k} \overline{W}^{(i)}/(k\nu/(1-\nu)),\ 2\leq k\leq 5.$$

Figure 5b shows this ratio for two through twenty-five stages in series for $\rho=1$. These results were obtained by extending recursive formulae (1), (2), and (3). In this extreme case of correlation, adding stages has an effect on system performance which depends markedly on the utilization rate; e.g., for $\nu=0.7$ system performance is improved through the first four stages and then is reduced. A utilization of 0.9 gives rise to much improved system performance through twenty-five stages.

CASES B and C: Two stage queues in series, finite (including zero) interstage storage.

For these cases the utilization is effectively reduced in value [17]. The maximum effective utilization is $\nu_{\max}=(q+1)/(q+2)$ where the queue in stage two is limited to a length of $q-1$ units. We consider the cases $q=1$ and $q=2$.

Figure 6 shows the mean waiting time at the first stage for $q=1$ and several values of ν. For this case each customer's waiting time at the second stage is simply the service time there so we are concerned only with the waiting time process at stage one. Figure 7 shows, for stage one, the ratio of mean waiting time at stage one with $\rho\neq 0$ to the mean waiting time at stage one with $\rho=0$.

Steady state results for the mean number of customers in the system, L, for $\rho=0$, $q=1$ and with utilization ν are given in Morse [15]; we have that

$$L=4\nu(2-\nu^2)/((2+\nu)(2-3\nu)). \tag{14}$$

For $\nu=0.4$, 0.5, and 0.6 and for $\rho=0$, the observed (expected) values of L are 1.55(1.53), 2.87(2.80), and 7.80(7.57) respectively. The observed values are from Figure 6.

For $\rho\neq 0$, again we see a dramatic effect in system performance. System performance deteriorates as the correlation ρ increases through positive values and improves as ρ decreases through negative values. Hence when there is no storage allowed before stage two, the departure from in-

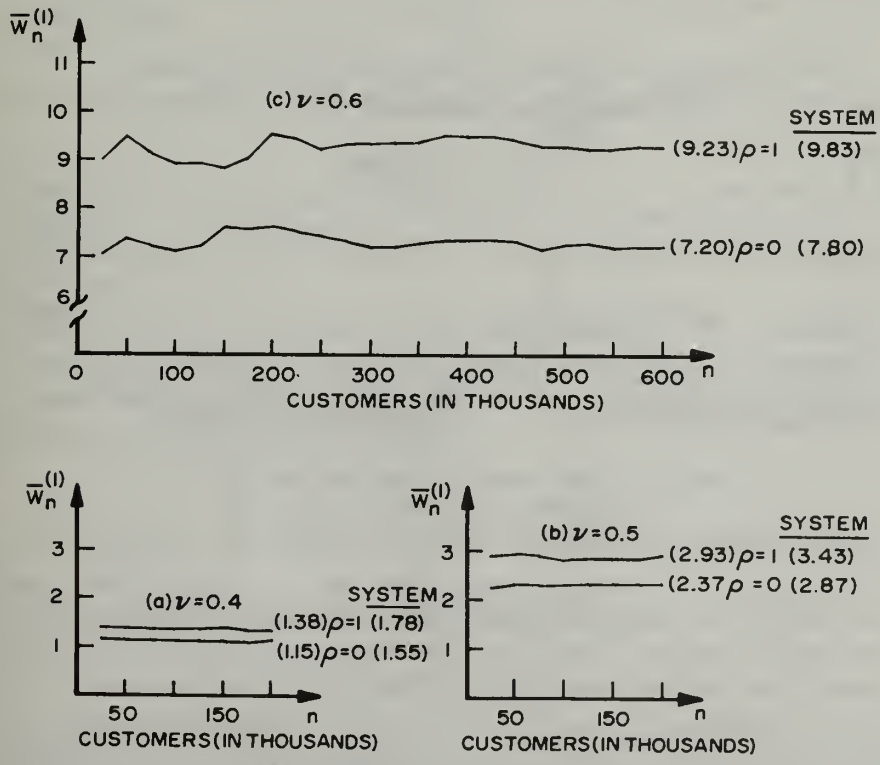


FIGURE 6.—Mean waiting time at first stage, 2Q, $q=1$.

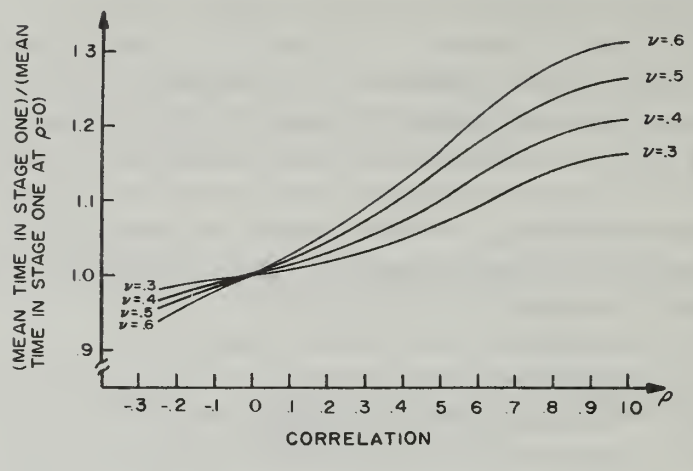


FIGURE 7.—Ratio of mean waiting times at first stage, 2Q, $q=1$.

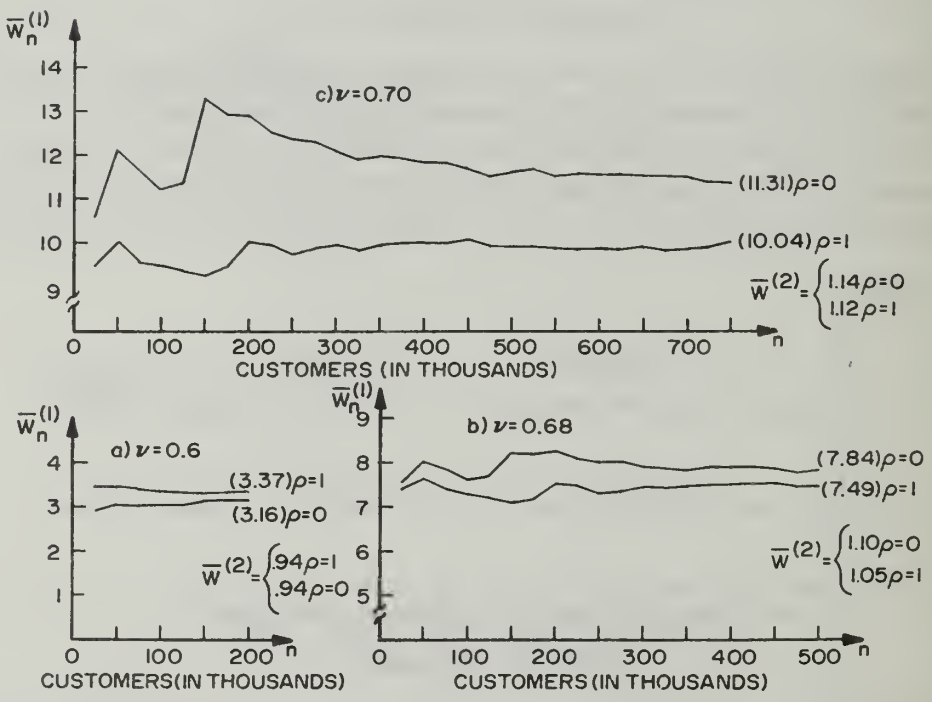


FIGURE 8.—Mean waiting time at first stage, 2Q, $q=2$.

dependence results in significantly different steady state behaviors, especially as the value ν_{max} is approached.

Finally, we consider the case $q=2$. Figure 8 shows the mean waiting time as a function of n at the first stage and the ending value for the mean waiting time in the second stage. The mean waiting time at stage two was very stable for all values of n so those values will not be illustrated. The effect for $\nu=0.6$ is in the same direction as for $q=1$ but reverses as ν increases so that for values close to ν_{max} the change in system performance is consistent with the $q=\infty$ case; that is, improvement for $\rho>0$ and deterioration for $\rho<0$. Figure 9 shows the effect in this case for $\rho\neq 0$.

5. SPECTRAL ANALYSIS OF $\{W_n^{(1)}\}$ AND $\{W_n^{(2)}\}$

In this section we analyze the time series $\{W_n^{(1)}\}$ and $\{W_n^{(2)}\}$ and apply a nonparametric test to the ratio of certain estimated power spectra associated with them.

Let $\{z_t,\ t=1, 2, \ldots, N\}$ be a realization of a stochastic process with mean μ and autocovariances γ_k, $k=1, 2, \ldots$. A study of a time series in terms of its autocovariances is referred to as a time domain analysis. Another type of analysis is concerned with the frequency content of the time series, namely spectral analysis [1, 8, 10]. The Fourier cosine transform of the autocovariances $\gamma_0, \gamma_1, \gamma_2, \ldots$, is called the power spectrum. Denoting the power spectrum by $f(\omega)$, we can write

$$f(\omega)=\frac{1}{\pi}\left[\gamma_0+2\sum_{k=1}^{\infty}\gamma_k \cos 2\pi\omega k\right],\ 0\leq\omega\leq 1/2 \tag{15}$$

and inverting $f(\omega)$ we can express γ_k as

$$\gamma_k=\int_0^{1/2} f(\omega)\cos 2\pi\omega k d\omega,\ k=0, 1, 2, \ldots \tag{16}$$

with γ_k estimated by

$$c_k=\frac{1}{N}\sum_{t=1}^{N-k}(z_{t+k}-\bar{z})(z_t-\bar{z}),$$

where

$$\bar{z}=\frac{1}{N}\sum_{t=1}^{N} z_t.$$

When $k=0$ we obtain the variance γ_0 of the process as the integral of the power spectrum:

$$\gamma_0=\int_0^{1/2} f(\omega)d\omega. \tag{17}$$

Thus the power spectrum can be considered as a decomposition of the variance at different frequencies.

To get sample results which are statistically consistent we do not estimate the spectrum at a particular frequency but instead estimate the average power about the frequency of concern. The average power corresponds to weighting the autocovariances in the time domain and we typically estimate $f(\omega)$ with the truncated estimate

$$\hat{f}(\omega_j)=\frac{1}{\pi}\left[\lambda_0 c_0+2\sum_{k=1}^{m}\lambda_k c_k \cos 2\pi\omega_j k\right] \tag{18}$$

where $\omega_j=j/(2m)$, $j=0, 1, 2, \ldots, m$ and the weights λ_k, $k=0, 1, 2, \ldots, m$, form a so-called lag window. We choose the Blackman-Tukey "hamming" window,

$$\lambda_k=0.54+0.46 \cos \pi k/m,\ k=0, 1, 2, \ldots, m \tag{19}$$

In (18), the sample autocovariances $c_{m+1}, c_{m+2}, \ldots$, are omitted since, for m sufficiently large, they should contribute little information. As a result, only m autocovariances need be calculated and savings in computation may be considerable. Considerable care must be used when selecting

m, however, because too large a value will increase the variance of the estimates and too small a value will not give enough resolution.

Next we examine several sample power spectra associated with the simulated waiting times for the two-server infinite interstage storage case. We take the simulated values $\{W_n^{(i)}\}$,

$$n=1, 2, \ldots, N; i=1, 2,$$

to be time series where, as before, $W_n^{(i)}$ is the total waiting time, queueing plus service, of customer n at server i. Figure 10 shows a portion of the sample spectra for $\{W_n^{(1)}\}$ and $\{W_n^{(2)}\}$,

$$n=1, 2, \ldots, 2000,$$

and for correlation values of $\rho=0$, 0.25, 0.50, and 1.0. Utilization, ν, is 0.90. The 2000 sample values were chosen from the end of a simulation run of length 30,000 to ensure that any possible effects of start-up conditions were eliminated. After making several pilot runs, m in equation (18) was set equal to 400. For $\rho=0.50$ and $\rho=1.0$ in Figure 10 it is obvious that the waiting times at the second server give rise to different spectra than the waiting times at the first server.

Since the integral of the power spectrum measures the variance of the process and the area under the sample spectrum should be indicative of the sample variance, we see that the effect of positive correlation is to *reduce* the variance of the waiting time process. Again this is consistent with Conolly's results [4] for the single server system in which a customer's service is completely determined by the length of the interarrival interval separating himself and his predecessor. For a utilization of 0.9, Conolly's system reduces the steady state variance of the waiting time from 81, for the classic M/M/1 system, to 1.16; the sample variance associated with the waiting times at server 2 in Figure 10d for $\rho=1.0$ is 2.05 (the sample variance associated with the waiting times at server 1 is 60.1). Recall from Section 4 that the condition $\rho=1.0$ for the correlation between a customer's service times at the two servers is equivalent to a correlation of ν, or 0.9 in this case, between his interarrival time and service time at the second server. Therefore, the reduction in variance is consistent with Conolly's results since the corresponding correlation in his system is one.

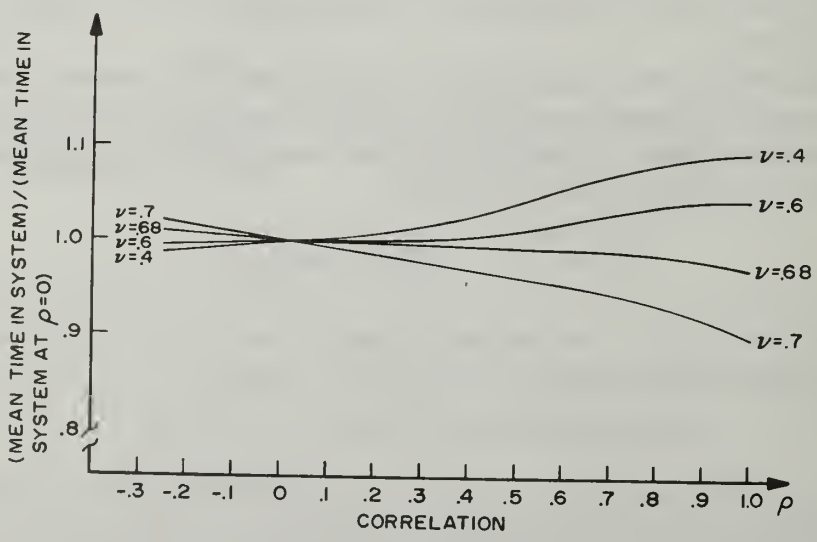


FIGURE 9.—Ratio of mean waiting times in system, 2Q, $q=2$.

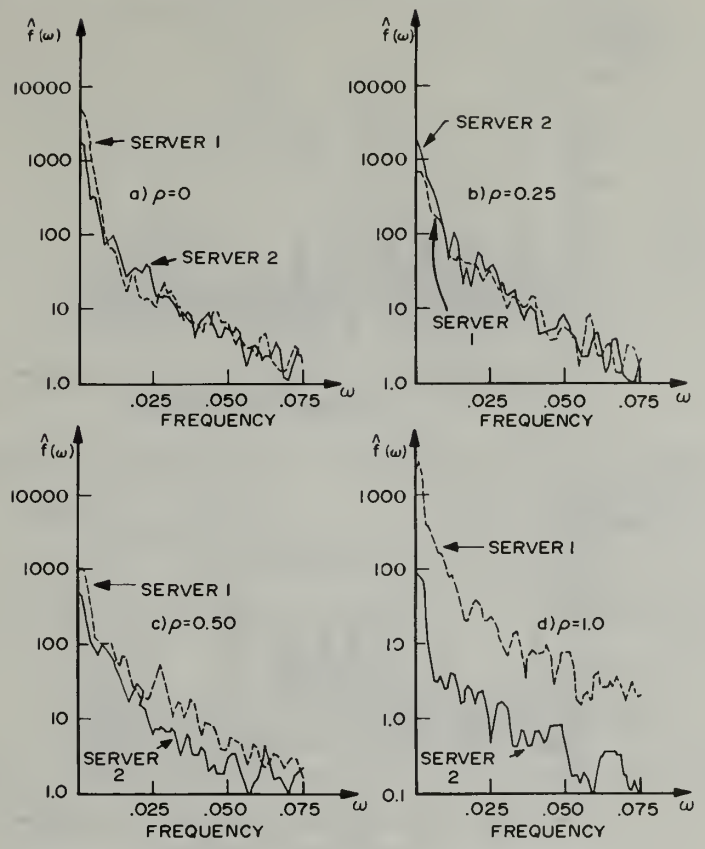


FIGURE 10.—A portion of waiting time spectra, 2Q, $\nu=0.9$, $q=\infty$.

Next we develop a nonparametric test for the hypothesis that $f^{(1)}(\omega)=f^{(2)}(\omega)$, $0\leq\omega\leq 0.5$, where $f^{(i)}(\omega)$ represents the power spectrum at frequency ω associated with the time series $\{W_n^{(i)}\}$, $n=1, 2, \ldots, N$; $i=1, 2$. The Blackman-Tukey "hamming" lag window in (19) gives rise to spectral estimates which are not independent, and so we employ the notion of equivalent independent estimates [10] which implies, for this window, that estimates are approximately independent if they are about 5/(4m) cycles apart. Since the estimates in (18) are separated by a basis frequency of 1/(2m) cycles, this spacing of 5/(4m) cycles amounts to taking, as independent, those estimates which are separated by an interval of 2.5 times the basic frequency. Since the waiting times are not normally distributed and the assumption of normality is implicit in the development of equivalent independent estimates, we take this spacing of 2.5 times the basic frequency simply to be a rough guide. Actually, the normality assumption is more critical for making distributional assumptions about the spectral estimates than for the usage here. To select a practical spacing and to reduce any possible effects of the normality assumption we take estimates at the frequencies $j/(2m)$, $j=1, 4, 7, \ldots$, to be approximately independent (the spacing here is 3 times the basic frequency). Therefore, of the 401 estimates in each spectrum partially illustrated in Figure 10, we take 134 estimates at the frequencies $j/800$, $j=1, 4, 7, \ldots, 399$, to be approximately independent.

Now for each approximately independent estimate we can regard the ratio $\hat{f}^{(1)}(\omega)/\hat{f}^{(2)}(\omega)$ as a Bernoulli trial (greater than one or less than one) and under the null hypothesis of homogeneity of the two spectra we can take as a test statistic the number of ratios which are less than unity. Figure

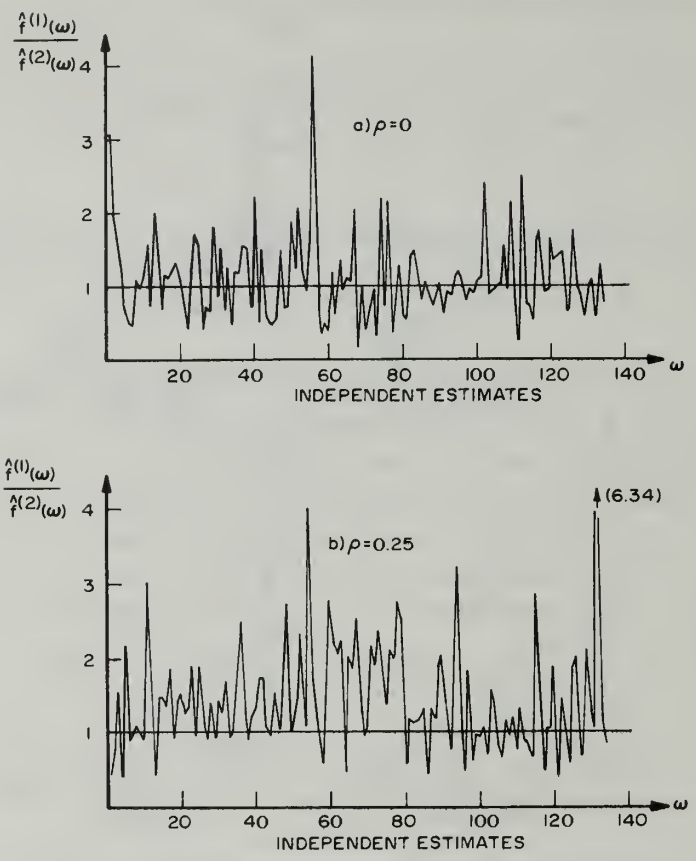


FIGURE 11.—Ratio of sample power spectral estimates, 2Q, $\nu=0.9$, $q=\infty$.

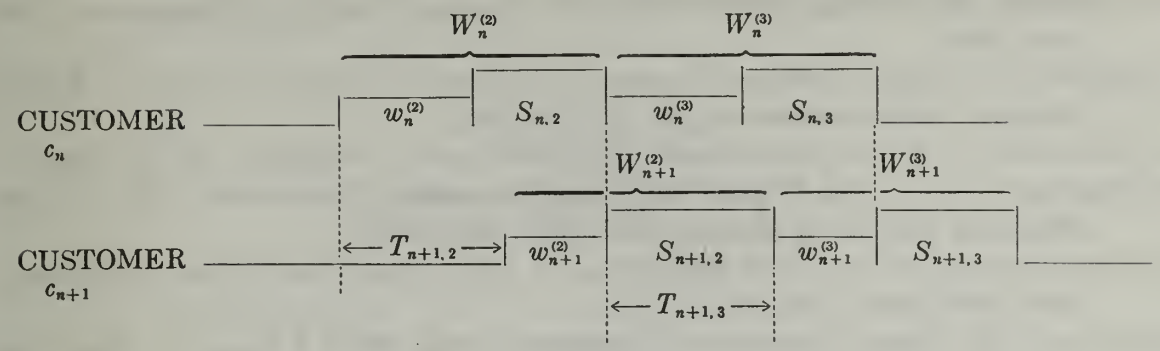
11 shows the ratio for $\rho=0$ and $\rho=0.25$. Of the 134 approximately independent ratios in Figure 11a, 64 are less than unity and in Figure 11b, 43 of the 134 ratios are less than unity. Under the null hypothesis, a ratio greater than unity is as equally likely as a ratio less than unity, and taking a normal approximation to the implied binomial distribution, we have a probability of 0.31 associated with observing 64 or fewer ratios less than unity in 11a and a probability of .001 associated with observing 43 or fewer ratios less than unity in 11b. Although not illustrated, the results for $\rho=0.50$ and $\rho=1.0$ are even more conclusive: for $\rho=0.50$, 23 of the 134 approximately independent estimates are less than unity, and for $\rho=1.0$, none of the ratios is less than unity. Therefore, for the case presented here we have good statistical evidence that the power spectra associated with the waiting times at each server are not homogeneous for correlation $\rho>0$. We expect similar results for other values of correlation, utilization, and interstage storage to obtain.

APPENDIX

In this appendix we show how a set of recursive formulae for the waiting times can be constructed for any number of queues in series where interstage storage is unlimited.

Referring to the following diagram for c_n and c_{n+1}'s queueing and service times at stages two and three (a continuation of the diagram preceding equation (1) in the text), we see that c_{n+1}'s interarrival time at stage three is

$$T_{n+1,3}=\begin{cases}T_{n+1,2}-W_n^{(2)}+S_{n+1,2} & \text{if } T_{n+1,2}\geq W_n^{(2)}\\ S_{n+1,2}, & \text{if } T_{n+1,2}<W_n^{(2)}.\end{cases}$$



Similar to equation (3) in the text, c_{n+1}'s waiting time at the third stage is

$$W_{n+1}^{(3)}=\begin{cases}S_{n+1,3}, & \text{if } T_{n+1,3}\geq W_n^{(3)}\\ W_n^{(3)}-T_{n+1,3}+S_{n+1,3}, & \text{if } T_{n+1,3}<W_n^{(3)}.\end{cases}$$

Comparing $T_{n+1,2}$ and $T_{n+1,3}$ we have, in general, for c_{n+1}'s interarrival time at stage i, $i=2, 3, \ldots,$

$$T_{n+1,i}=\begin{cases}T_{n+1,i-1}-W_n^{(i-1)}+S_{n+1,i-1}, & \text{if } T_{n+1,i-1}\geq W_n^{(i-1)}\\ S_{n+1,i-1}, & \text{if } T_{n+1,i-1}<W_n^{(i-1)}.\end{cases}$$

Similarly, comparing $W_{n+1}^{(1)}$, $W_{n+1}^{(2)}$ and $W_{n+1}^{(3)}$ gives a general recursive formula for c_{n+1}'s waiting time at stage i, $i=1, 2, \ldots,$

$$W_{n+1}^{(i)}=\begin{cases}S_{n+1,i}, & \text{if } T_{n+1,i}\geq W_n^{(i)}\\ W_n^{(i)}-T_{n+1,i}+S_{n+1,i}, & \text{if } T_{n+1,i}<W_n^{(i)}.\end{cases}$$

Thus, we can obtain the recursive formulae for any number of queues in series.

REFERENCES

[1] Anderson, T. W., *The Statistical Analysis of Time Series* (John Wiley & Sons, Inc., New York, 1971).

[2] Bhat, U. N., "Queueing Systems with First-Order Dependence," Opsearch, *6*, 1–24 (1969).

[3] Burke, P. J., "The Output of a Queueing System," Operations Research, *4*, 699–704 (1956).

[4] Conolly, B. W., "The Waiting Time Process for a Certain Correlated Queue," Operations Research, *16*, 1006–1015 (1968).

[5] Conolly, B. W. and N. Hadidi, "A Correlated Queue," Journal of Applied Probability, *6*, 122–136 (1969).

[6] Conway, R. W., "Some Tactical Problems in Digital Simulation," Management Science, *16*, 47–61 (1963).

[7] Downton, F., "Bivariate Exponential Distributions in Reliability Theory," J. Royal Statist. Society, Series B, *32*, 408–417 (1970).

[8] Fishman, G. S. and P. J. Kiviat, "The Analysis of Simulation-Generated Time Series," Management Science, *13*, 525–557 (1967).

[9] Jackson, R. R. P., "Queueing Systems with Phase Type Service," Operations Research, *5*, 109–120 (1954).

[10] Jenkins, G. M., "General Considerations in the Analysis of Spectra," Technometrics, *3*, 133–166 (1961).

[11] Kesten, H., "Random Difference Equations and Renewal Theory for Products of Random Matrices," Acta Mathematica, *131*, 207–248 (1973).

[12] Kibble, W. F., "A Two-Variate Gamma Type Distribution," Sankhyā, *5*, 137–150 (1941).

[13] Kohberger, R. C., "On Certain Multivariate Exponential Distributions," Ph. D. Thesis, Rensselaer Polytechnic Institute, Troy, New York (1975).

[14] Krishnamoorthy, A. S. and M. Parthasarty, "A Multivariate Gamma Type Distribution," Ann. Math. Statist., *22*, 549–557 (1951).

[15] Morse, P. M., *Queues, Inventories and Maintenance* (John Wiley & Sons, Inc., New York, 1958).

[16] Paulson, A. S., "A Characterization of the Exponential Distribution and a Bivariate Exponential Distribution," Sankhyā, Series A, *35*, 69–78 (1973).

[17] Saaty, T. L., *Elements of Queueing Theory with Applications* (McGraw-Hill: New York, 1961).

[18] Wicksell, S. D., "On Correlation Functions of Type III," Biometrika, *25*, 121–133 (1933).

SINGLE-LANE BRIDGE SERVING TWO-LANE TRAFFIC

Z. Eshcoli and I. Adiri

Technion—Israel Institute of Technology
Haifa, Israel

ABSTRACT

This paper presents a mathematical model of a single-lane bridge serving two-way traffic in alternating directions (with an FIFO rule observed within each directional queue). While the bridge serves cars moving in one direction, cars approaching from the opposite direction wait in a queue at its foot. When cars in the current direction finish crossing the bridge, it begins serving cars from the other direction, if any are present. A newly-arrived car finding an empty bridge mounts it immediately. Several cars moving in the same direction may occupy the bridge simultaneously. The crossing speed is assumed to be constant, and the arrival processes in both directions are assumed to be independent, homogeneous Poisson processes. A generalization of the alternating-priority models [1, 2] is developed to arrive at the Laplace-Stieltjes transform and the expected value of the flow time (the time interval between the moments of arrival at the bridge and departure from it) for steady state conditions. The results are discussed and some examples are presented graphically.

1. INTRODUCTION

Cars arrive at a single-lane bridge in two opposite directions, numbered 1 and 2, according to independent homogeneous Poisson processes with arrival rates λ_1 and λ_2, respectively. A car going in direction i ($i=1, 2$), hereinafter referred to as "type-i car," may cross the bridge if the latter is empty or carrying cars of the same type; otherwise (i.e., if cars of the other type are crossing) it must wait in a queue at the foot of the bridge.

The time interval between the moment when the bridge starts serving type-i cars and the moment when it carries none of them is called "type-i phase." At the end of type-i phase ($i=1, 2$), either the bridge is empty or type-j cars ($j \neq i$) have queued up at its foot. This queue, called the initial queue, now mounts the bridge in order of arrival. The crossing speed is assumed constant, hence the crossing time is also constant. This assumption is not unrealistic since usually the speed limit on such a narrow bridge is far below the free speed of modern cars.

When there is no queue of type-i cars but the bridge still carries them, a newly-arrived type-i car does not have to wait, thus spending only the crossing time in the system.

The crucial difference between the above described and an alternating-priorities queueing model ([1, 2, Chapter 9]) is that here the service facility (the bridge) can accommodate several

customers (cars) simultaneously; accordingly, the term "service" will need a special definition. given later on.

Our aim is to find the steady-state distribution and the expectation of the flow time (the time interval between the moments of arrival at the bridge and departure from it) of a car in the system as a function of the bridge's length (or equivalently the time spent crossing the bridge). Although our analysis is based on [1] and [2], these models may be derived as special cases of the model discussed in this paper. The situation described above is not limited to the case of a narrow bridge, but is applicable to any one-lane road servicing two directional traffic, a common situation when repairing a road. Assume that a two-lane road of given length has to be repaired. The repair will be done by choosing one lane for traffic alternatively. The maintenance chief would like to do the work in one stretch to save set-up costs; on the other hand, this policy creates long queues and in turn high flow times. Having the necessary cost data, the results obtained may serve as a guideline in determining the optimal partition for the repair of the road.

At the end a discussion of the model is given, and the expected steady-state flow time as a function of the length of the bridge, for specific parameters, is presented graphically.

Related models were studied by several authors. Darroch, Newell and Morris [3] considered a model in which a vehicle-actuated traffic light controls two intersecting traffic streams. The light is kept green for lane i ($i=1, 2$) until any existing queue of type-i cars has been discharged (the "discharging time"), and further until a headway of duration at least β_i is detected in the subsequent arrivals (the "extension time"). The main difference between [3] and our model is that in [3] the discharging time and the extension time are independent random variables. Thus the light may change to green for lane-i even if there are no cars present in this lane, and this light stays green only for the extension time during which cars arriving from the intersecting lane have to wait even though no type-i cars are present. This case cannot happen in our model where a busy bridge is available for type-i cars iff type-i cars are crossing it. In another paper, Hawkes [4] assumed generally distributed crossing times and alternating priorities discipline. The expected waiting time of a type-i car was calculated and this result, obviously, coincides with the result in [1] and subsequently may be obtained as a special case of our model. Tanner [5] discussed a similar model in which the crossing times were also constants but the queue discipline was different: A type-i car could cross the bridge if there were no type-j cars ($j\neq i$) on the bridge and the last type-i car has started crossing at least β_i time units ago. An explicit formula for the expected waiting time of a type-i car was presented only when $\beta_1=0$ or $\beta_2=0$.

2. MATHEMATICAL MODEL

2.1. Basic Relations

Let Y_i ($i=1, 2$) be the mounting time* of a type-i car with the exception of a type-i car initiating a type-i phase which has Y°_i as its mounting time, and let S_i be the constant crossing time of a type-i car. The mounting times are assumed to be non-negative arbitrarily-distributed random variables independent of each other and of the interarrival times, and possessing finite second moments. The crossing times are assumed to be finite positive constants. Let the arrival

*Defined as the time interval between the moments two successive cars (present in the system) begin to cross the bridge.

process of cars in direction i ($i=1, 2$) be a homogeneous Poisson with average of λ_i cars per time unit.

We denote:

(1) $$\lambda=\lambda_1+\lambda_2,$$

and

(2) $$\rho_i=\lambda_i E(Y_i) \qquad i=1, 2,$$

then under the above assumptions the system is non-saturated if:

(3) $$\rho=\rho_1+\rho_2<1,$$

see Section 4.

At steady-state the system undergoes cycles of length T_c, the components of each cycle being an idle period T_a (the time period during which the bridge is empty), and a busy period T_b (the time interval between two successive idle periods). Thus:

(4) $$T_c=T_a+T.$$

Two types of busy periods are observed: T_{b_1} and T_{b_2} where $T_{bj}(j=1, 2)$ starts with the arrival of a type-j car to an empty bridge and terminates when the system is empty. Hence:

(5) $$E(T_c)=E(T_a)+E(T_b)=\frac{1}{\lambda}+\frac{\lambda_1}{\lambda}E(T_{b_1})+\frac{\lambda_2}{\lambda}E(T_{b_2}).$$

Following the notation in [2] T_{bj} may be subdivided into subcycles of successive flows in alternating directions denoted by T_{kj}.

(6) $$T_{bj}=\sum_{k=1}^{\infty} T_{kj}, \qquad j=1, 2.$$

Every subcycle T_{kj} (except when $k=j=1$) comprises two successive phases of flow and counter-flow as follows:

(7)
$$\begin{aligned} T_{11}&=T_{111} \\ T_{k1}&=T_{2,k-1,1}+T_{1k1}, \qquad k>1 \\ T_{k2}&=T_{2k2}+T_{1k2}, \qquad k\geq 1 \end{aligned}$$

where T_{ikj} is the k-th phase of flow of type-i cars in a type-j busy period.
Thus, the two types of busy periods may be described as follows:

Type-1

Phases: $T_{111} \mid T_{211} \mid T_{121} \mid T_{221} \mid T_{131} \mid \cdots \mid T_{2,k-1,1} \mid T_{1k1} \mid \cdots$

Subcycles: $T_{11} \mid T_{21} \mid T_{31} \mid \quad \mid T_{k1} \mid$

$\longleftarrow T_{b_1} \longrightarrow$

Type-2

Phases: $T_{212} \mid T_{112} \mid T_{222} \mid T_{122} \mid \cdots \mid T_{2k2} \mid T_{1k2} \mid \cdots$

Subcycles: $T_{12} \mid T_{22} \mid \mid T_{k2} \mid$

$\leftarrow\!\!-\!\!-\!\!-

2.2. Length of Phase

Consider T_{1k1}, $k>1$. Clearly $T^{(2)}_{1k1}=0$ whenever $T^{(1)}_{1k1}=0$ so that for $k>1$ we have:

$$T_{1k1}=\begin{cases} T^{(1)}_{1k1}+T^{(2)}_{1k1}, & T^{(1)}_{1k1}>0 \\ 0, & T^{(1)}_{1k1}=0; \end{cases} \tag{11}$$

where $T^{(1)}_{1k1}$ and $T^{(2)}_{1k1}$ are defined and discussed by equation (8) and its sequel.

If the first component, $T^{(1)}_{1k1}$, is positive it can be treated as a busy period in an M/G/1 model where the service time is Y_1, and the service time of the "first customer" being the time required by the initial queue of type-1 cars, formed during $T_{2,k-1,1}$ to mount the bridge. Therefore.

$$T^{(1)}_{1k1}=\sum_{i=1}^{\infty} t_i, \tag{12}$$

where t_1 is the mounting time required by the initial queue and t_i ($i>1$) is the time required by the cars arrived during t_{i-1} to mount the bridge.

By well-known results ([2, p. 151]):

$$L_{T^{(1)}_{1k1}}(z)=L_{t_1}(z+\lambda_1-\lambda_1 L_1(z)),\ k>1 \tag{13}$$

where

$$L_1(z)=L_{Y_1}(z+\lambda_1-\lambda_1 L_1(z)), \tag{14}$$

Let $N_{2,k-1,1}$ be the number of type-1 cars arrived during $T_{2,k-1,1}$ then:

$$E(e^{-zt_1}|T_{2,k-1,1}=t,\ N_{2,k-1,1}=n)=\begin{cases} L_{Y^{\circ}_1}(z)[L_{Y_1}(z)]^{n-1}, & n>0 \\ 1, & n=0 \end{cases} \tag{15}$$

Withdrawing the conditions in (15) we have:

$$L_{t_1}(z)=L_{T_{2,k-1,1}}(\lambda_1)+\frac{L_{Y^{\circ}_1}(z)}{L_{Y_1}(z)}\,(L_{T_{2,k-1,1}}(\lambda_1-\lambda_1 L_{Y_1}(z))-L_{T_{2,k-1,1}}(\lambda_1)),\ k>1. \tag{16}$$

Substituting (16) in (13) we obtain:

$$L_{T^{(1)}_{1k1}}(z)=L_{T_{2,k-1,1}}(\lambda_1)+\frac{L_{Y^{\circ}_1}(z+\lambda_1-\lambda_1 L_1(z))}{L_1(z)}\cdot(L_{T_{2,k-1,1}}(\lambda_1-\lambda_1 L_1(z))-L_{T_{2,k-1,1}}(\lambda_1)),\qquad k>1 \tag{17}$$

where $L_1(z)$ is given by (14).

As for the second subphase in T_{1k1}: let

$$\overline{T}^{(2)}_{1k1}\equiv T^{(2)}_{1k1}/T^{(1)}_{1k1}>0$$

and let $M^{(2)}_{1k1}$ be the number of type-1 cars arrived during $\overline{T}^{(2)}_{1k1}$. Denote by τ_{m1}, $m\geq 1$, the time elapsed between the moments of arrival of the $(m$-1)-th and the m-th car in $\overline{T}^{(2)}_{1k1}$. τ_{m1} are independent r.v. distributed as τ_1. $\tau_1<S_1$ because otherwise $\overline{T}^{(2)}_{1k1}$ would terminate. Hence the density function of τ_1 takes the form:

$$f_{\tau_1}(t)=\frac{\lambda_1 e^{-\lambda_1 t}}{1-e^{-\lambda_1 S_1}},\qquad 0<t<S_1. \tag{18}$$

A type-1 car arriving at the system during $\overline{T}^{(2)}_{1k1}$ could cross the bridge iff the preceding type-1 car were still on it—an event which occurs with probability $1-e^{-\lambda_1 S_1}$, hence:

$$P(M^{(2)}_{1k1}=n)=(1-e^{-\lambda_1 S_1})^n e^{-\lambda_1 S_1}, \qquad n\geq 0; \qquad k>1. \tag{19}$$

Now for $k>1$,

$$\overline{T}^{(2)}_{1k1}=\begin{cases} S_1+\sum_{m=1}^{M^{(2)}_{1k1}} \tau_{m1}, & M^{(2)}_{1k1}>0 \\ S_1, & M^{(2)}_{1k1}=0. \end{cases} \tag{20}$$

Expressing (20) in terms of L.S. transform we arrive at:

$$L_{\overline{T}^{(2)}_{1k1}}(z)=\frac{\lambda_1+z}{\lambda_1+ze^{(\lambda_1+z)S_1}}, \qquad k>1. \tag{21}$$

When $k=1$ it is true that:

$$T_{111}=T^{(2)}_{111}=S_1+\sum_{m=1}^{M^{(2)}_{111}-1} \tau_{m1}, \tag{22}$$

but

$$P(M^{(2)}_{111}=n)=(1-e^{-\lambda_1 S_1})^{n-1}e^{-\lambda_1 S_1}, \; n=1, 2, \ldots, \tag{23}$$

(It is clear from the definition that at least one type-1 car appears during T_{111}.) Therefore:

$$L_{T_{111}}(z)=\frac{\lambda_1+z}{\lambda_1+ze^{(\lambda_1+z)S_1}}. \tag{24}$$

When $k>1$, $T_{1k1}=0$ iff in the preceding phase, $T_{2,k-1,1}$, there were no arrivals of type-1 cars. The probability of this event is:

$$P_{(T_{1k1}=0)}=\int_0^\infty e^{-\lambda_1 t}dF_{T_{2,k-1,1}}(t)=L_{T_{2,k-1,1}}(\lambda_1), \qquad k>1 \tag{25}$$

By (11), (25) and the independence of $T^{(1)}_{1k1}|T^{(1)}_{1k1}>0$ and $\overline{T}^{(2)}_{1k1}$:

$$E(e^{-zT_{1k1}})=L_{T_{2,k-1,1}}(\lambda_1)+(1-L_{T_{2,k-1,1}}(\lambda_1))L_{\overline{T}^{(2)}_{1k1}}(z)E\left(e^{-zT^{(1)}_{1k1}}|T^{(1)}_{1k1}>0\right), \; k>1. \tag{26}$$

But

$$E\left(e^{-zT^{(1)}_{1k1}}|T^{(1)}_{1k1}>0\right)=\int_0^\infty e^{-zt}dF_{T^{(1)}_{1k1}|T^{(1)}_{1k1}>0}(t)=\int_{0+}^\infty e^{-zt}d\left\{\frac{F_{T^{(1)}_{1k1}}(t)}{1-P_{(T_{1k1}=0)}}\right\}$$

$$=\frac{1}{1-L_{T_{2,k-1,1}}(\lambda_1)}\left\{L_{T^{(1)}_{1k1}}(z)-L_{T_{2,k-1,1}}(\lambda_1)\right\}, \; k>1. \tag{27}$$

Substituting (27) in (26) yields:

$$L_{T_{1k1}}(z)=L_{T_{2,k-1,1}}(\lambda_1)+L_{\overline{T}^{(2)}_{1k1}}(z)\left(L_{T^{(1)}_{1k1}}(z)-L_{T_{2,k-1,1}}(\lambda_1)\right), \; k>1. \tag{28}$$

From (17), (21) and (28) we finally obtain:

$$L_{T_{1k1}}(z)=L_{T_{2,k-1,1}}(\lambda_1)+\frac{\lambda_1+z}{\lambda_1+ze^{(\lambda_1+z)S_1}}\frac{L_{Y^0_1}(z+\lambda_1-\lambda_1 L_1(z))}{L_1(z)}\cdot\left(L_{T_{2,k-1,1}}(\lambda_1-\lambda_1 L_1(z))-L_{T_{2,k-1,1}}(\lambda_1)\right), \; k>1, \tag{29}$$

where $L_1(z)=L_{Y_1}(z+\lambda_1-\lambda_1 L_1(z))$.

Similarly, by symmetry, we have:

$$(30)\quad L_{T_{2k1}}(z)=L_{T_{1k1}}(\lambda_2)+\frac{\lambda_2+z}{\lambda_2+ze^{(\lambda_1+z)S_2}}\frac{L_{Y^{\circ}_2}(z+\lambda_2-\lambda_2L_2(z))}{L_2(z)}\cdot(L_{T_{1k1}}(\lambda_2-\lambda_2L_2(z))-L_{T_{1k1}}(\lambda_2)),\ k\geq 1,$$

where

$$L_2(z)=L_{Y_2}(z+\lambda_2-\lambda_2L_2(z)).$$

Theoretically, $L_{T_{ik_i}}(z)$ is obtainable recursively from the above relations: $L_{T_{111}}(z)$ is given by (24); substituting $L_{T_{111}}(z)$ in (30) for $k=1$ yields $L_{T_{211}}(z)$; substituting $L_{T_{211}}(z)$ in (29) for $k=2$ yields $L_{T_{121}}(z)$, etc. $L_{T_{i\,k+1}}(z)$ are obtained by change of indices in $L_{T_{ik1}}(z)$.

Differentiating (24), (29) and (30) with respect to z at $z=0$, yields the expected value of the length of a phase:

$$(31)\quad E(T_{111})=\frac{e^{\lambda_1S_1}-1}{\lambda_1}$$

$$(32)\quad E(T_{1k1})=\frac{\rho_1}{1-\rho_1}E(T_{2,\,k-1,\,1})+\left(\frac{E(Y^{\circ}_1-Y_1)}{1-\rho_1}+\frac{e^{\lambda_1S_1}-1}{\lambda_1}\right)(1-L_{T_{2,\,k-1,\,1}}(\lambda_1)),\ k>1$$

$$(33)\quad E(T_{2k1})=\frac{\rho_2}{1-\rho_2}E(T_{1k1})+\left(\frac{E(Y^{\circ}_2-Y_2)}{1-\rho_2}+\frac{e^{\lambda_2S_2}-1}{\lambda_2}\right)(1-L_{T_{1k1}}(\lambda_2)),\ k\geq 1.$$

Solving the set (31), (32), and (33), we have:

$$(34)\quad E(T_{1k1})=E(T_{111})r^{k-1}+\left(\frac{E(Y^{\circ}_1-Y_1)}{1-\rho_1}+\frac{e^{\lambda_1S_1}-1}{\lambda_1}\right)\sum_{m=1}^{k-1}r^{k-1-m}(1-L_{T_{1m1}}(\lambda_1))$$
$$+\frac{\rho_1}{1-\rho_1}\left(\frac{E(Y^{\circ}_2-Y_2)}{1-\rho_2}+\frac{e^{\lambda_2S_2}-1}{\lambda_2}\right)\sum_{m=1}^{k-1}r^{k-1-m}(1-L_{T_{1m1}}(\lambda_2)),\ k>1,$$

and

$$(35)\quad E(T_{2k1})=\frac{\rho_2}{1-\rho_2}E(T_{111})r^{k-1}+\frac{\rho_2}{1-\rho_2}\left(\frac{E(Y^{\circ}_1-Y_1)}{1-\rho_1}+\frac{e^{\lambda_1S_1}-1}{\lambda_1}\right)\sum_{m=1}^{k-1}r^{k-1-m}(1-L_{T_{2m1}}(\lambda_1))$$
$$+\left(\frac{E(Y^{\circ}_2-Y_2)}{1-\rho_2}+\frac{e^{\lambda_2S_2}-1}{\lambda_2}\right)\sum_{m=1}^{k}r^{k-m}(1-L_{T_{1m1}}(\lambda_2)),\ k\geq 1,$$

where

$$r=\frac{\rho_1\rho_2}{(1-\rho_1)(1-\rho_2)},$$

and $E(T_{111})$ is given by (31).

Substituting (31), (34) and (35) in (7) yields the expected length of the subcycles $E(T_{k1})$, $k\geq 1$. Thus, to find the Laplace-Stieltjes transform of the flow time of a type-i car, it is left to find $E(T_c)$ and $L_{U_{ki}}(z)$, $i=1, 2$; $k=1, 2, \ldots$, In the next section we calculate $E(T_c)$.

2.3. Expected Length of Cycle

Let N_{ij} $(i, j=1, 2)$ be the number of type-i phases in type-j busy period. The distributions and expectations of these r.v. were found in [2, p. 199]:

$$(36)\quad P(N_{11}=1)=P(T_{121}=0)=\int_0^{\infty}e^{-\lambda_1 t}dF_{T_{111}}(t)=L_{T_{111}}(\lambda_1),$$

and

(37)

$$P(N_{11}=k)=P(T_{1k1}>0,\ T_{1,\,k+1,\,1}=0)=P(T_{1,\,k+1,\,1}=0)-P(T_{1k1}=0)=L_{T_{2k1}}(\lambda_1)-L_{T_{2,\,k-1,\,1}}(\lambda_1),\ k>1.$$

Hence:

$$E(N_{11})=1+\sum_{k=1}^{\infty}(1-L_{T_{2k1}}(\lambda_1)) \tag{38}$$

and similarly:

$$E(N_{21})=\sum_{k=1}^{\infty}(1-L_{T_{1k1}}(\lambda_2)). \tag{39}$$

The expected values of N_{12} and N_{22} are obtained from (38) and (39) by change of indices.

From (31), (32), (33), (38) and (39), we have:

$$\sum_{k=1}^{\infty}E(T_{1k1})=\frac{1}{1-r}\left\{\frac{e^{\lambda_1 S_1}-1}{\lambda_1}+\left(\frac{E(Y^{\circ}_1-Y_1)}{1-\rho_1}+\frac{e^{\lambda_1 S_1}-1}{\lambda_1}\right)\cdot E(N_{11}-1)\right. \tag{40}$$
$$\left.+\frac{\rho_1}{1-\rho_1}\left(\frac{E(Y^{\circ}_2-Y_2)}{1-\rho_2}+\frac{e^{\lambda_2 S_2}-1}{\lambda_2}\right)E(N_{21})\right\}$$

and

$$\sum_{k=1}^{\infty}E(T_{2k1})=\frac{1}{1-r}\left\{\frac{\rho_2}{1-\rho_2}\,\frac{e^{\lambda_1 S_1}-1}{\lambda_1}+\frac{\rho_2}{1-\rho_2}\left(\frac{E(Y^{\circ}_1-Y_1)}{1-\rho_1}+\frac{e^{\lambda_1 S_1}-1}{\lambda_1}\right)\cdot E(N_{11}-1)\right. \tag{41}$$
$$\left.+\left(\frac{E(Y^{\circ}_2-Y_2)}{1-\rho_2}+\frac{e^{\lambda_2 S_2}-1}{\lambda_2}\right)E(N_{21})\right\}.$$

Hence:

$$E(T_{b1})=\sum_{k=1}^{\infty}[E(T_{1k1})+E(T_{2k1})]=\frac{1}{1-\rho}\left\{\left(\frac{E(Y^{\circ}_1-Y_1)}{1-\rho_1}+\frac{e^{\lambda_1 S_1}-1}{\lambda_1}\right)E(N_{11})\right. \tag{42}$$
$$\left.+E(Y^{\circ}_1-Y_1)+(1-\rho_2)\left(\frac{E(Y^{\circ}_2-Y_2)}{1-\rho_2}+\frac{e^{\lambda_2 S_2}-1}{\lambda_2}\right)E(N_{21})\right\}.$$

$E(T_{b_2})$ is obtained from (42) by change of indices.

Finally from (42) and (5):

$$E(T_c)=\frac{1}{\lambda}+\frac{1}{\lambda(1-\rho)}\left\{\lambda_1(1-\rho_1)\left(\frac{E(Y^{\circ}_1-Y_1)}{1-\rho_1}+\frac{e^{\lambda_1 S_1}-1}{\lambda_1}\right)E(N_{11})+\lambda_1 E(Y_1-Y^{\circ}_1)\right. \tag{43}$$
$$+\lambda_2(1-\rho_1)\left(\frac{E(Y^{\circ}_1-Y_1)}{1-\rho_1}+\frac{e^{\lambda_1 S_1}-1}{\lambda_1}\right)E(N_{12})+\lambda_1(1-\rho_2)\left(\frac{E(Y^{\circ}_2-Y_2)}{1-\rho_2}+\frac{e^{\lambda_2 S_2}-1}{\lambda_2}\right)$$
$$\left.\cdot E(N_{21})+\lambda_2(1-\rho_2)\left(\frac{E(Y^{\circ}_2-Y_2)}{1-\rho_2}+\frac{e^{\lambda_2 S_2}-1}{\lambda_2}\right)E(N_{22})+\lambda_2 E(Y_2-Y^{\circ}_2)\right\}$$

where $E(N_{ij})$, $i, j=1, 2$, are given by equations (38) and (39).

Substituting $E(T_c)$ and the previous results in $L_{W_1}(z)$ (equation (10)), it still remains to find the L.S. transform of the flow time of a type-1 car in a subcycle. This is done in the next section.

2.4. Flow Time

Consider

$$T_{k1}=T_{2,\,k-1,\,1}+T_{1k1}=T_{2,\,k-1,\,1}+T^{(1)}_{1k1}+T^{(2)}_{1k1},\ k>1.$$

The flow time of a type-1 car in $T^{(2)}_{1k1}$ is S_1. Let U'_{k1} $k>1$ be the flow time of a type-1 car in

$$T\equiv T_{2,\,k-1,\,1}+T^{(1)}_{1k1} \tag{44}$$

T may be treated as a delayed busy period in an M/G/1 model where $T_{2,k-1,1}$ is the delay interval and $T^{(1)}_{1k1}$ is treated as in (12). Using known results for the flow time in an M/G/1 model where the busy period is delayed and the first customer has a different service time ([2, p. 153]):

$$L_{U'_{k1}}(z)=\frac{e^{-S_1z}\{L_{Y_1}(z)(1-L_{T_{2,\,k-1,\,1}}(\lambda_1))-L_{Y^\circ_1}(z)(L_{T_{2,\,k-1,\,1}}(z)-L_{T_{2,\,k-1,\,1}}(\lambda_1))\}}{E(T_{2,\,k-1,\,1}+T^{(1)}_{1k1})(\lambda_1L_{Y_1}(z)-\lambda_1+z)},\qquad k>1. \tag{45}$$

By (11) and (25) we have:

$$L_{U_{k1}}(z)=\frac{E(T_{2,\,k-1,\,1}+T^{(1)}_{1k1})}{E(T_{k1})}\,L_{U'_{k1}}(z)+(1-L_{T_{2,\,k-1,\,1}}(\lambda_1))\,\frac{E(T^{(2)}_{1k1})}{E(T_{k1})}\,e^{-S_1z},\qquad k>1. \tag{46}$$

Substituting (45) in (46) yields:

$$L_{U_{k1}}(z)=\frac{e^{-S_1z}}{E(T_{k1})}\left\{\frac{L_{Y_1}(z)(1-L_{T_{2,\,k-1,\,1}}(\lambda_1))-L_{Y^\circ_1}(z)(L_{T_{2,\,k-1,\,1}}(z)-L_{T_{2,\,k-1,\,1}}(\lambda_1))}{\lambda_1L_{Y_1}(z)-\lambda_1+z}\right.$$
$$\left.+\frac{e^{\lambda_1S_1}-1}{\lambda_1}(1-L_{T_{2,\,k-1,\,1}}(\lambda_1))\right\},\ k>1 \tag{47}$$

and similarly:

$$L_{U_{k2}}(z)-\frac{e^{-S_1z}}{E(T_{k2})}\left\{\frac{L_{Y_1}(z)(1-L_{T_{2k2}}(\lambda_1))-L_{Y^\circ_1}(z)(L_{T_{2k2}}(z)-L_{T_{2k2}}(\lambda_1))}{\lambda_1L_{Y_1}(z)-\lambda_1+z}+\frac{e^{\lambda_1S_1}-1}{\lambda_1}(1-L_{T_{2k2}}(\lambda_1))\right\},\ k\geqq 1. \tag{48}$$

By definition:

$$L_{U_{11}}(z)=e^{-S_1z}. \tag{49}$$

Substituting (47), (48) and (49) in (10), we finally obtain the Laplace transform of the flow time for a type-1 car:

$$L_{W_1}(z)=\frac{e^{-S_1z}}{\lambda E(T_c)}\left\{1+\frac{e^{\lambda_1S_1}-1}{\lambda_1}(\lambda_1E(N_{11})+\lambda_2E(N_{12}))+\frac{1}{\lambda_1L_{Y_1}(z)-\lambda_1+z}\sum_{k=1}^{\infty}[L_{Y_1}(z)(\lambda_1(1-L_{T_{2k1}}(\lambda_1))\right.$$
$$\left.+\lambda_2(1-L_{T_{2k2}}(\lambda_1)))-L_{Y^\circ_1}(z)(\lambda_1(L_{T_{2k1}}(z)-L_{T_{2k1}}(\lambda_1))+\lambda_2(L_{T_{2k2}}(z)-L_{T_{2k2}}(\lambda_1)))]\right\} \tag{50}$$

Differentiating (50) with respect to z at $z=0$ we obtain the expected flow time of a type-1 car:

$$E(W_1)=S_1+\frac{1}{\lambda E(Tc)\cdot 2(1-\rho_1)^2}\left\{(1-\rho_1)\sum_{k=1}^{\infty}[\lambda_1E(T^2_{2k1})+\lambda_2E(T^2_{2k2})]+(\lambda_1E(Y_1{}^2)\right.$$
$$+2(1-\rho_1)E(Y^\circ_1))\sum_{k=1}^{\infty}[\lambda_1E(T_{2k1})+\lambda_2E(T_{2k2})]+(\lambda_1E(Y_1{}^2)E(Y^\circ_1-Y_1)$$
$$\left.+(1-\rho_1)E(Y^\circ_1)^2-Y_1{}^2))(\lambda_1E(N_{11}-1)+\lambda_2E(N_{12}))\right\} \tag{51}$$

Now

$$\sum_{k=1}^{\infty} E(T_{2k1})$$

is given by (41);

$$\sum_{k=1}^{\infty} E(T_{2k2})$$

is obtained from (40) by change of indices; differentiating (30) twice with respect to z at $z=0$, we have $E(T_{2k1}^2)$; differentiating (29) twice with respect to z at $z=0$ and changing indices, we have $E(T_{2k2}^2)$. These results are to be substituted in (51) to yield the expected value of the flow time of a type-1 car.

3. SPECIAL CASES

3.1 $F_{Y^\circ_i}(y) \equiv F_{Y_i}(y)$, $i=1, 2$.

We assumed that the first type-i car ($i=1, 2$) in each subcycle has a different mounting time. This is usually the case in real life. However, assuming that the mounting times of all type-i cars are identically distributed simplifies the equations and may provide adequate approximations.

Hence replacing Y°_i by Y_i ($i=1, 2$) in (40) and (41) and changing indices in (40), we have:

$$\sum_{k=1}^{\infty} E(T_{2k1}) = \frac{(1-\rho_1)(1-\rho_2)}{1-\rho} \left\{ \frac{\rho_2}{1-\rho_2} \frac{e^{\lambda_1 S_1}-1}{\lambda_1} E(N_{11}) + \frac{e^{\lambda_2 S_2}-1}{\lambda_2} E(N_{21}) \right\} \tag{52}$$

$$\sum_{k=1}^{\infty} E(T_{2k2}) = \frac{(1-\rho_1)(1-\rho_2)}{1-\rho} \left\{ \frac{e^{\lambda_2 S_2}-1}{\lambda_2} E(N_{22}) + \frac{\rho_2}{1-\rho_2} \frac{e^{\lambda_1 S_1}-1}{\lambda_1} E(N_{12}) \right\}. \tag{53}$$

Differentiating (24), (29) and (30) twice with respect to z at $z=0$ we obtain difference equations for the second moments of the phase length.

Summing these equations we obtain two independent equations for

$$\sum_{k=1}^{\infty} E(T_{1k1}^2)$$

and

$$\sum_{k=1}^{\infty} E(T_{2k1}^2)$$

whose solution is:

$$\sum_{k=1}^{\infty} E(T_{1k1}^2) = \frac{1}{1-r^2} \left\{ \left(\frac{\lambda_1 E(Y_1^2)}{(1-\rho_1)^3} + \frac{2\rho_1}{1-\rho_1} \frac{e^{\lambda_1 S_1}-1}{\lambda_1} \right) \sum_{k=1}^{\infty} E(T_{2k1}) + \frac{2e^{\lambda_1 S_1}(e^{\lambda_1 S_1}-1-\lambda_1 S_1)}{\lambda_1^2} E(N_{11}) \right\} \tag{54}$$

$$+ \frac{1}{1-r^2} \left(\frac{\rho_1}{1-\rho_1} \right)^2 \left\{ \left(\frac{\lambda_2 E(Y_2^2)}{(1-\rho_2)^3} + \frac{2\rho_2}{1-\rho_2} \frac{e^{\lambda_2 S_2}-1}{\lambda_2} \right) \sum_{k=1}^{\infty} E(T_{1k1}) + \frac{2e^{\lambda_2 S_2}(e^{\lambda_2 S_2}-1-\lambda_2 S_2)}{\lambda_2^2} E(N_{21}) \right\}$$

(55)

$$\sum_{k=1}^{\infty} E(T_{2k1}^2) = \frac{1}{1-r^2} \left\{ \left(\frac{\lambda_2 E(Y_2^2)}{(1-\rho_2)^3} + \frac{2\rho_2}{1-\rho_2} \frac{e^{\lambda_2 S_2}-1}{\lambda_2} \right) \sum_{k=1}^{\infty} E(T_{1k1}) + \frac{2e^{\lambda_2 S_2}(e^{\lambda_2 S_2}-1-\lambda_2 S_2)}{\lambda_2^2} E(N_{21}) \right\}$$

$$+ \frac{1}{1-r^2} \left(\frac{\rho_2}{1-\rho_2} \right)^2 \cdot \left\{ \left(\frac{\lambda_1 E(Y_1^2)}{(1-\rho_1)^3} + \frac{2\rho_1}{1-\rho_2} \frac{e^{\lambda_1 S_1}-1}{\lambda_1} \right) \sum_{k=1}^{\infty} E(T_{2k1}) + \frac{2e^{\lambda_1 S_1}(e^{\lambda_1 S_1}-1-\lambda_1 S_1)}{\lambda_1^2} E(N_{11}) \right\}.$$

Replacing Y°_i by Y_i $(i=1, 2)$ in (51), we obtain:

$$(56)\quad E(W_1)=S_1+\frac{1}{\lambda E(T_c)}\frac{1}{2(1-\rho_1)^2}\Big\{(1-\rho_1)\sum_{k=1}^{\infty}(\lambda_1 E(T_{2k1}^2)+\lambda_2 E(T_{2k2}^2))$$
$$+(\lambda_1 E(Y_1^2)+2(1-\rho_1)E(Y_1))\sum_{k=1}^{\infty}(\lambda_1 E(T_{2k1})+\lambda_2 E(T_{2k2}))\Big\}.$$

Now

$$\sum_{k=1}^{\infty}E(T_{2k1}^2)$$

is given by (55);

$$\sum_{k=1}^{\infty}E(T_{2k2}^2)$$

is obtained from (54) by change of indices;

$$\sum_{k=1}^{\infty}E(T_{2k1})$$

and

$$\sum_{n=1}^{\infty}E(T_{2k2})$$

are given in (52) and (53). Substituting these values in (56), we have the expected flow time of a type-1 car in terms of the known parameters of the system. $E(W_2)$ is obtained from (56) by change of indices, and $E(W)$ is obtained from (9).

3.2. Alternating Priorities

If the crossing times are negligible then the problem of simultaneous service disappears and the only service that cars receive at the bridge is mounting it. Hence substituting $S_i=0$, $i=1, 2$, in the above discussion we obtain the known formulas for the L.S. transform and the expected value of the flow time under alternating-priorities rule ([1, 2]). Furthermore, the case of alternating-priorities with set-up times may also be derived if the first service in each phase, Y°_i, $i=1, 2$, is decomposed into a sum of the (independent) set-up time and the "ordinary" service time.

4. DISCUSSION

4.1. Non-saturation Conditions

The distribution of $T^{(2)}_{1k1}$ $k>1$ is independent of k (equation (21)). In the same manner it can be proved that $T^{(2)}_{ikj}$, $k>1$, is independent of k and j.

Define:

$$(57)\qquad \overline{Y}^{\circ}_1=Y^{\circ}_1+T^{(2)}_{221},\ \overline{Y}^{\circ}_2=Y^{\circ}_2+T^{(2)}_{121}$$

With this definition, ignoring the first and last phases in each busy period, our model becomes identical, as far as saturation is concerned, to an alternating priorities model in which the service time for the i-th priority class $(i=1, 2)$ is distributed as Y_i, except for the first customer in a phase whose service time is distributed as $\overline{Y}^{\circ}_i$. (Obviously, the first and last phases in each busy period and the different service time of the first customer do not affect saturation.) Hence the non-satura-

tion condition for our model is the same as that for the above described alternating priorities model ([1, 2, Chapter 9]), namely the condition stated in (3).

4.2. Mounting Times

When T_{ikj} $k>1$ starts the queue of type-i cars formed at the foot of the bridge during the previous phase mounts the bridge. The mounting time of a car was defined as the time elapsed between the moments when two consecutive cars present in the system begin to cross the bridge. Clearly, the mounting time is the time needed to pass a distance equal to the length of a car and a minimal safety distance between two consecutive cars. (In the subcycles $T^{(2)}_{ikj}i$, $j=1, 2$, $k\geq 1$ the minimal safety distance is not necessarily kept.) The mounting time of the first car in a phase does not share the same distribution because the first car has some additional preparations to make before mounting the bridge and here no overlapping activities are possible.

4.3. Graphical Representation

It is assumed, for simplicity, that the two priority classes differ only in their arrival rates, i.e., they have the same crossing time S, and their mounting time is distributed as Y including the first car in a phase.

Furthermore let us denote:

A=bridge's length

ϕ=crossing velocity (constant)

c=car's length (assumed to be uniformly distributed on the interval (3, 5))

D=safety distance

ψ=mounting velocity (constant)

Distances are measured in meters and the unit of time is a minute. The crossing velocity was assumed to be constant, as a direct consequence the crossing time $S=A/\phi$ is constant too. Assume further that ψ is constant, then in view of 4.2, Y must be proportional to

$$\frac{c+D}{\psi}$$

and in fact, equality was assumed. Figure 1 shows the behaviour of the expected flow times $E(W_1)$ and $E(W_2)$ as a function of the bridge's length—A, with $\lambda_1=5$ cars/min., $\lambda_2=10$ cars/min., D=3 meters, $\psi=150$ meters/min., $\phi=500$ meters/min., $E(Y)=0.47$ min., $\rho=.7$.

Since $\lambda_1<\lambda_2$, it takes more time to clear the bridge of type-2 cars, thus accounting for the greater expected flow time of type-1 cars.

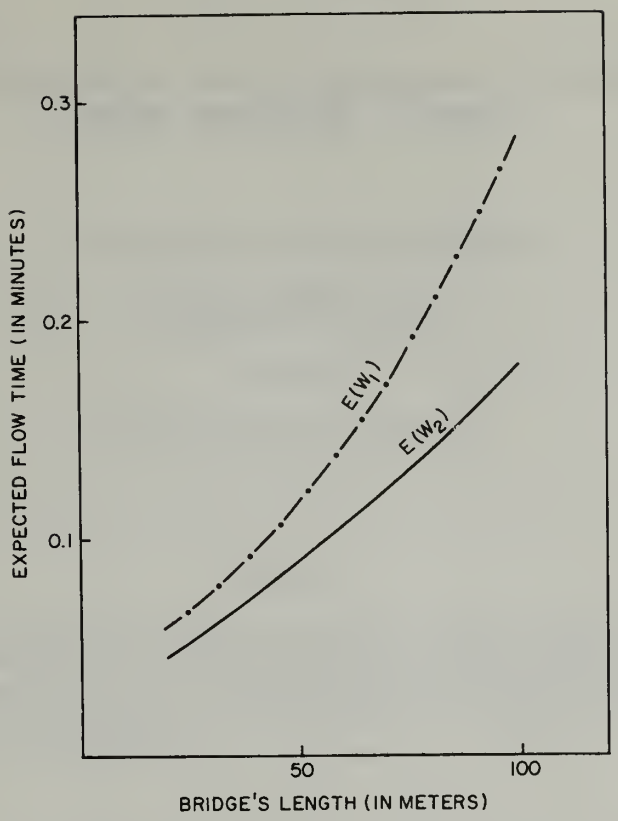


FIGURE 1.—Expected flow time as a function of the bridge's length.

REFERENCES

[1] Avi-Itzhak, B., W. L. Maxwell and L. W. Miller, "Queueing with Alternating Priorities," Operations Research, *13*, 306–318 (1965).

[2] Conway, R. W., W. L. Maxwell and L. W. Miller, *Theory of Scheduling*, (Addison-Wesley, 1967).

[3] Darroch, J. N., G. F. Newell and R. W. J. Morris, "Queueing for a Vehicle-Actuated Traffic Light," Operations Research, *12*, 882–895 (1964).

[4] Hawkes, A. G., "Queueing at Traffic Intersections," Proceedings, Second Symposium on the Theory of Traffic Flow, London (1963).

[5] Tanner, J. C., "A Problem of Interference Between Two Queues," Biometrika, *40*, 58–69 (1953).

OPTIMAL CONTROL FOR MULTI-SERVERS QUEUEING SYSTEMS UNDER PERIODIC REVIEW*

C. C. Huang,† S. L. Brumelle and K. Sawaki

University of British Columbia
Vancouver, B.C., Canada

I. Vertinsky

International Institute of Management
Berlin, Germany

ABSTRACT

This paper deals with the problem of finding the optimal dynamic operating policy for an M/M/S queue. The system is observed periodically, and at the beginning of each period the system controller selects the number of service units to be kept open during that period. The optimality criterion used is the total discounted cost over a finite horizon.

INTRODUCTION

Most of the related studies reported recently in the literature focused upon controls of one-server queues over infinite horizon (Heyman [3], Bell [2], Sobel [5], Balachandran [1]).

Zacks and Yadin [6] dealt with the case of finite horizon for M/M/1 system with variable service intensity under non periodic review. In their model decision epochs occur immediately after changes in queue size. They identify an "optimal" policy resorting to a conjecture and imposing restrictive assumptions. Magazine [4] studied M/M/S queues with finite waiting capacity under periodical review and convex increasing holding costs over finite horizons.

All the results in Magazine [4] are derived on the basis of (1) a misspecified holding cost function which ignores the fact that customers may be turned away when waiting room capacity is saturated, and (2) an incorrect argument that the distribution of number served by the i^{th} server in an m-server system is identical to the distribution in an n-server system where $n \neq m$.

This paper considers a similar model to the one outlined by Magazine with two major modifications:

(1) Waiting room capacity is taken to be infinite.

(2) The cost structure is generalized to permit different holding costs in different periods.

---

*The research reported in this paper was supported in part by the Defence Research Board of Canada and The International Institute of Management, Berlin.

†Currently affiliated with Memorial University of Newfoundland, St. John's, Newfoundland, Canada.

MODEL FORMULATION

System Structure

(i) Assume there are s servers in parallel.

(ii) Assume a Poisson arrival stream.

(iii) Assume that the decision points are at equally spaced time intervals. Without loss of generality, we assume that these points are 0, 1, 2, ...

(iv) Assume that the service times have independent and identical exponential distributions and are independent of the arrival process.

Cost Structure

Let $S(l|k)$ denote the cost of changing from k to l open servers at a decision epoch, and then operating the l servers until the next decision epoch. We assume that $S(l|k)$ has the following properties:

(i) $S(l|k)$ is convex in l for each k;

(ii) For fixed j and l with $l \leq j$, the function $S(j|k) - S(l|k)$ is nonincreasing in k, and equal to

$$S(j|k-1) - S(l|k-1) \text{ for } k \leq l \text{ or } k-1 \geq j.$$

For example, we might take

$$S(l|k) = (l-k)^+A + (k-l)^+B + lC$$

where A is the cost of opening a closed server, B is the cost of closing an open server, C is the cost of operating an open server for one period, and

$$k^+ = \max\ (k, 0).$$

It is easy to check that $S(l|k)$ of this form satisfies (i) and (ii).

In addition to the above switching and operating costs, we include a customer holding cost. Let $K_n(i)$ denote the cost incurred during period n if it ends with i customers in the system.

Markov Decision Structure

Define $\phi_n(i, \alpha, k)$ as the minimum expected cost with i customers present, k servers open and n periods remaining in the horizon, using discount factor α, $0 \leq \alpha \leq 1$. We take the length of time between decision epochs to be the unit of time. The recursive relationships are

$$\phi_n(i, \alpha, k) = \min_{0 \leq l \leq s} \psi_n(l|i, \alpha, k) \qquad n = 1, 2, \cdots, N$$

where

$$\psi_n(l|i, \alpha, k) = S(l|k) + \sum_{j=0}^{\infty} P_{ij}{}^{l}[K_n(j) + \alpha\phi_{n-1}(j, \alpha, l)],\ \phi_0(i, \alpha, k) = 0,$$

Z_n is the number of customers in the system at the beginning of period n, and

$$P_{ij}{}^{l} = [Z_n = j | Z_{n+1} = i \cap l \text{ servers are open}].$$

Policy Structure

For each period index n and $k=0, 1, 2, \cdots, s$ define

$$I_n(l|k)=\{i:\phi_n(i, \alpha, k)=\psi_n(l|i, \alpha, k)\},\ l=0, 1, 2, \cdots, s.$$

Define a policy π by

$$\pi(n, i, k)=l \text{ if } i\epsilon 1_n(l|k)$$

with the interpretation: If there are k servers open and i customers present at the nth decision epoch and $i\epsilon I(l|k)$, then change the number of servers open to l.

The optimality of the above policy is obvious, since for any n, k and i, one can evaluate $\psi_n(l|i, \alpha, k)$ for $l=0, 1, \cdots, s$ and assign the i to the appropriate set $I_n(l|k)$ which by definition will yield the minimum expected cost.

It is reasonable to expect that for fixed n and k, the $I_n(l|k)$ are disjoint intervals. In this case, the policy can be specified by a nondecreasing sequence of control limits

$$i_0^*(k, n)\leq i_1^*(k, n), \cdots, \leq i_s^*(k, n)\leq i_{s+1}^*(k, n)$$

with the interpretation that if there are k servers open with i customers present at a decision epoch with n periods remaining and $i_l^*(k, n)\leq i<i_{l+1}^*(k, n)$, then change the number of servers open to l. In this case we define the policy π by

$$\pi(n, i, k)=l \text{ if } i_l^*(k, n)\leq i<i_{l+1}^*(k, n) \tag{1}$$

The limits $i^*{}_l(k, n)$ can be obtained recursively from the set functions $1_n(l|k)$ as follows: set

$$i_{s+1}^*(k, n)=+\infty;$$

and for $j=s, s-1, \ldots, 0$, set

$$i_j^*(k, n)=i_{j+1}^*(k, n) \text{ if } I_n(j|k)=\phi$$

and otherwise set $i_j^*(k, n)=\min\{i: i\epsilon I_n(j|k\}$.

A similar policy was suggested by Magazine [4].

In the next section of the paper we show that an opitmal policy of the above control limit type exists.

The Control Limit Form of the Optimal Policy

Let Z_n and L_n respectively denote the number of customers in the system and the number of servers on at the decision epoch with n periods remaining. It is convenient to model the $\{Z_n, L_n\}$ process in the following way. Define

$$X_n=\{\tau_{n1}, \tau_{n2}, \tau_{n3}, \ldots; S_{n1}, S_{n2}, S_{n3}, \ldots,\}$$

to be a sequence of independent random variables with each τ having the exponential interarrival distribution and each S having the exponential service time distribution.

If

$$r=\max\left[i:\sum_{j=1}^{i}\tau_{nj}<1\right],$$

then r customers will arrive during period n at times

$$\tau_{n1},\ \tau_{n1}+\tau_{n2},\ \cdots,\ \tau_{n1}+\tau_{n2}+\cdots+\tau_{nr}$$

measured from the beginning of the period. The services times are $S_{n1}, S_{n2}, \ldots, S_{nr}$ for the r customers who arrive during the period and times

$$S_{n,r+1},\ S_{n,r+2},\ \cdots,\ S_{n\,r+Z_n}$$

for the Z_n customers in the system at the beginning of the period.

We also assume that $\{X_n: n=1, 2, 3, \cdots,\}$ is a sequence of independent vectors. Then Z_{n-1} is a function of Z_n, L_n, and X_n for a given policy π, i.e. for each policy π, there is a function f such that $Z_{n-1}=f(Z_n, L_n, X_n)$ $n=1, 2, \cdots,$.

We assume that during period n, customers have priority in the order that the service times are listed in the vector X_n. Thus the customers who arrive during the period have priority over those customers present at the beginning of the period. The priorities are preemptive, and customers who are preempted resume their service at the point it was preempted.

Thus the process representing the number of customers in the system is modelled somewhat differently than usual, since we choose new service times for all customers in the system at a decision epoch, even if they are in service. In addition, the queue discipline is not the usual one. However, since the service times are exponential, the probability distribution of the stochastic process $\{Z_n\}$ will be the same as in a normal M/M/S system. That is P_{ij}^{l} is the same for our system as for the usual first-come first-serve M/M/S system.

Consequently, any result that depends only on P_{ij}^{l} will be true in both our system and the usual system. In particular, all results except for Lemma 1 hold for the usual M/M/S system.

In the following discussion, the period index n will be clear from the context and will be dropped for notational convenience.

Let $N_l(X|i)$ be the random variable denoting Z_{n-1} given that $\pi(Z_n, L_n)=l$ and $Z_n=i$; that is $N_l(x|i)$ is the number of customers left at the end of the period, given that i customers are present at the beginning of the period, that l servers are open during the period, and that $X_n=x$.

The following discussion uses the difference notion

$$\frac{\Delta f(y)}{\Delta y}=\frac{f(y+\Delta y)-f(y)}{\Delta y}$$

LEMMA 1: For any realization x of X, $N_m(x|i+u)-N_{m'}(x|i)$ is non-negative and nondecreasing in i for $m'\geq m$ and $u\geq 0$.

PROOF: For any realization, we have

$$\frac{\Delta N_m(x|i+u)}{\Delta i} \text{ and } \frac{\Delta N_{m'}(x|i)}{\Delta i}$$

are either 0 or 1.

To see this, note that the additional customer has the lowest priority, and so does not affect any of the other customers. The number left at the end of the period increases by one if the additional customer does not complete service, and remains the same if he does. It is equally clear that

$$N_{m'}(x|i+1)=N_{m'}(x|i)+1$$

implies

$$N_m(x|i+u+1)=N_m(x|i+u)+1.$$

This gives

$$\frac{\Delta N_m(x|i+u)}{\Delta i}\geq\frac{\Delta N_{m'}(x|i)}{\Delta i}.$$

LEMMA 2: Assume that

$$\frac{\Delta^2 K(i)}{\Delta i}\geq 0.$$

Then for $m'\geq m$,

$$\sum_{j=0}^{\infty}[P_{ij}^m-P_{ij}^{m'}]K(j)$$

is nondecreasing in i.

PROOF:

$$\sum_{j=0}^{\infty}[P_{ij}^m-P_{ij}^{m'}]K(j)=E[K(N_m(X|i))-K(N_{m'}(X|i))].$$

For $X=x$, we have

$$\begin{aligned}\frac{\Delta K(N_m(x|i))}{\Delta i}&=\frac{\Delta K(y)}{\Delta y}\Big|_{y=N_m(x|i)}\frac{\Delta N_m(x|i)}{\Delta i}\\&\geq\frac{\Delta K(y)}{\Delta y}\Big|_{y=N_m(x|i)}\frac{\Delta N_{m'}(x|i)}{\Delta i}\\&\geq\frac{\Delta K(y)}{\Delta y}\Big|_{y=N_{m'}(x|i)}\frac{\Delta N_{m'}(x|i)}{\Delta i}\\&=\frac{\Delta K(N_{m'}(x|i))}{\Delta i},\end{aligned}$$

where the first inequality follows by Lemma 1 and the second, since $N_m(x|i)\geq N_{m'}(x|i)$ and

$$\frac{\Delta^2 K(i)}{\Delta i^2}\geq 0.$$

Thus $K(N_m(x|i)-K(N_{m'}(x|i))$ is nondecreasing in i. Hence the expectation is also nondecreasing in i.

THEOREM 1: Assume that

$$\frac{\Delta^2 K_1(i)}{\Delta i^2}\geq 0.$$

Then for $n=1$ the optimal policy is of the control limit form given by display (1).

PROOF: Recall that

$$\phi_1(i,\alpha,k)=\min_{0\leq l\leq s}\psi_1(l|i,\alpha,k).$$

We prove the theorem by showing that $\psi_1(l|i, \alpha, k)-\psi_1(l+1|i, \alpha, k)$ is nondecreasing in i which imply that $\psi_1(l|i, \alpha, k)$ intersects $\psi_1(l+1|i, \alpha, k)$ at most once and from below.

The above difference can be written as

$$[S(l|k)-S(l+1|k)]+\sum_{j=0}^{\infty}(P_{ij}{}^{l}-P_{ij}^{l+1})K_1(j).$$

The first term is constant in i and the second term is nondecreasing in i by Lemma 2.

LEMMA 3: Assume that

$$\frac{\Delta^2 K_1(i)}{\Delta i^2}\geq 0.$$

Then

$$\frac{\Delta^2\phi_1(i, \alpha, k)}{\Delta i\Delta k}\leq 0.$$

PROOF: Let

$$\phi_1(i, \alpha, k)=S(k_1|k)+\sum_{j=0}^{\infty}P_{ij}{}^{k_1}K_1(j)$$

$$\phi_1(i, \alpha, k+1)=S(k_2|k+1)+\sum_{j=0}^{\infty}P_{ij}{}^{k_2}K_1(j)$$

$$\phi_1(i+1, \alpha, k)=S(k_1'|k)+\sum_{j=0}^{\infty}P_{i+1,j}^{k_1'}K_1(j)$$

$$\phi_1(i+1, \alpha, k+1)=S(k_2'|k+1)+\sum_{j=0}^{\infty}P_{i+1,j}^{k_2'}K_1(j)$$

First we show that $k_2\geq k_1$ and $k_2'\geq k_1'$. Since $\phi_1(i, \alpha, k)\leq\psi_1(l|i, \alpha, k)$ for each l, we have

$$S(k_1|k)-S(l|k)\leq\sum_{j=0}^{\infty}(P_{ij}{}^{l}-P_{ij}{}^{k_1})K_1(j).$$

If $l\leq k_1$, then by assumption (ii) of the cost structure

$$S(k_1|k+1)-S(l|k+1)\leq S(k_1|k)-S(l|k)$$

Consequently,

$$S(k_1|k+1)-S(l|k+1)\leq\sum_{j=0}^{\infty}(P_{ij}{}^{l}-P_{ij}{}^{k_1})K_1(j).$$

Regrouping terms provides

$$\psi_1(k_1|i, \alpha, k+1)\leq\psi_1(l|i, \alpha, k+1) \text{ for } l\leq k_1.$$

This expression demonstrates that $k_2\geq k_1$. A similar argument shows that $k_2'\geq k_1'$.
From Theorem 1 we also know that $k_1'\geq k_1$ and $k_2'\geq k_2$.
The proof proceeds by fixing k and considering three cases.

1. Suppose $i\geq i^*_{k+1}(k)$. By definition of $\phi_1(i, \alpha, k)$ we have

$$S(k_1|k)+\sum_{j=0}^{\infty}P_{ij}{}^{k_1}K_1(j)\leq S(l|k)+\sum_{j=0}^{\infty}P_{ij}{}^{l}K_1(j).$$

From Theorem 1, we know that $k_1 \geq k+1$. So the above inequality implies

$$S(k_1|k+1)+\sum_{j=0}^{\infty} P_{ij}^{k_1}K_1(j) \leq S(l|k+1)+\sum_{j=0}^{\infty} P_{ij}^{l}K_1(j)$$

for $l>k_1$, and thus $k_2 \leq k_1$. However, we showed earlier that $k_2 \geq k_1$. Therefore $k_1=k_2$. In this case,

$$\frac{\Delta}{\Delta k}\phi_1(i, j, k)=S(k_1|k+1)-S(k_1|k)$$

and

$$\frac{\Delta^2}{\Delta i \Delta k}\phi_1(i, j, k)=0.$$

2. Suppose $i<i^*_{k+1}(k+1)$. In this case a similar argument shows that again $k_1=k_2$ and

$$\frac{\Delta}{\Delta k}\phi_1(i, j, k)=S(k_2|k+1)-S(k_2|k).$$

Thus

$$\frac{\Delta^2}{\Delta i \Delta k}\phi_1(i, j, k)=0$$

3. Suppose $i^*_{k+1}(k+1) \leq i<i^*_{k+1}(k)$. From Theorem 1, we have that $k_1 \leq k$, $k'_1 \leq k$, $k_2>k$, and $k_2'>k$. Combining these inequalities with those verified at the first part of the proof provides $k_1 \leq k_1' \leq k+1<k_2 \leq k_2'$.

Since $k_2 \geq k_1'$

$$\phi_1(i, \alpha, k)-\phi_1(i, \alpha, k+1)=S(k_1|k)+\sum_{j=0}^{\infty} P_{ij}^{k_1}K_1(j)-S(k_2|k+1)-\sum_{j=0}^{\infty} P_{ij}^{k_2}K_1(j)$$

$$<S(k_1'|k)+\sum_{j=0}^{\infty} P_{ij}^{k_1'}K_1(j)-S(k_2|k+1)-\sum_{j=0}^{\infty} P_{ij}^{k_2}K_1(j) \leq S(k_1'|k)+\sum_{j=0}^{\infty} P_{i+1,j}^{k_1'}K_1(j)-S(k_2|k+1)$$

$$-\sum_{j=0}^{\infty} P_{i+1,j}^{k_2}K_1(j)<S(k_1'|k)+\sum_{j=0}^{\infty} P_{i+1,j}^{k_1'}K_1(j)-S(k_2'|k+1)-\sum_{j=0}^{\infty} P_{i+1,j}^{k_2'}K_1(j)$$

$$=\phi_1(i+1, \alpha, k)-\phi_1(i+1, \alpha, k+1),$$

where the middle inequality follows from Lemma 2. Thus

$$\frac{\Delta^2}{\Delta i \Delta k}\phi_1(i, \alpha, k) \leq 0$$

for $i^*_{k+1}(k+1) \leq i<i^*_{k+1}(k)$, which concludes the proof.

LEMMA 4: Assume that

$$\frac{\Delta^2[K_{n+1}(i)+\alpha\phi_n(i, \alpha, k)]}{\Delta i^2} \geq 0$$

for $n=0, 1$ and all $k=0, 1, \cdots, s$. Then for

$$m'>m, \sum_{j=0}^{\infty}\{P_{ij}^{m}[K_2(j)+\alpha\phi_1(j, \alpha, m)]-P_{ij}^{m'}[K_2(j)+\alpha\phi_1(j, \alpha, m')]\}$$

is nondecreasing in i.

PROOF:

$$\sum_{j=0}^{\infty}\{P_{ij}^{m}[K_2(j)+\alpha\phi_1(j,\alpha,m)]-P_{ij}^{m'}[K_2(j)+\alpha\phi_1(j,\alpha,m')]\}$$

$$=E\{[K_2(N_m(X|i))+\alpha\phi_1(N_m(X|i),\alpha,m)]-[K_2(N_{m'}(X|i))+\alpha\phi_1(N_{m'}(X|i),\alpha,m')]\}$$

For $X=x$ we have

$$\frac{\Delta[K_2(N_m(x|i))+\alpha\phi_1(N_m(x|i).\,\alpha,m)]}{\Delta i}=\left[\frac{\Delta K_2(y)}{\Delta y}\bigg|_{y=N_m(x|i)}+\alpha\frac{\Delta\phi_1(y,\alpha,m)}{\Delta y}\bigg|_{y=N_m(x|i)}\right]\frac{\Delta N_m(x|i)}{\Delta i}$$

$$\geq\left[\frac{\Delta K_2(y)}{\Delta y}\bigg|_{y=N_m(x|i)}+\alpha\frac{\Delta\phi_1(y,\alpha,m)}{\Delta y}\bigg|_{y=N_m(x|i)}\right]\frac{\Delta N_{m'}(x|i)}{\Delta i}$$

$$\geq\left[\frac{\Delta K_2(y)}{\Delta y}\bigg|_{y=N_{m'}(x|i)}+\alpha\frac{\Delta\phi_1(y,\alpha,m)}{\Delta y}\bigg|_{y=N_{m'}(x|i)}\right]\frac{\Delta N_{m'}(x|i)}{\Delta i}\geq\left[\frac{\Delta K_2(y)}{\Delta y}\bigg|_{y=N_{m'}(x|i)}\right.$$

$$\left.+\alpha\frac{\Delta\phi_1(y,\alpha,m')}{\Delta y}\bigg|_{y=N_{m'}(x|i)}\right]\frac{\Delta N_{m'}(x|i)}{\Delta i}\equiv\frac{\Delta[K_2(N_{m'}(x|i))+\alpha\phi_1(N_{m'}(x|i),\alpha,m')]}{\Delta i},$$

where the first inequality is by Lemma 1, the second since

$$N_m(x|i)\geq N_{m'}(x|i)\text{ and }\frac{\Delta^2[K_2(i)+\alpha\phi_1(i,\alpha,m)]}{\Delta i^2}\geq 0$$

and the third by Lemma 3.
Consequently,

$$[K_2(N_m(x|i))+\alpha\phi_1(N_m(x|i),\alpha,m)]-[K_2(N_{m'}(x|i))+\alpha\phi_1(N_{m'}(x|i),\alpha,m')]$$

is nondecreasing in i for every realization of X. Hence its expectation is also nondecreasing in i and the lemma is proved.

Thus, by using Lemma 4 and following the same arguments as in Theorem 1, we have ϕ_2 has an optimal control limit policy given by display (1).
By using the above arguments recursively, the following theorem is obtained.

THEOREM 2: Assume that

$$\frac{\Delta^2}{\Delta i^2}[K_{n+1}(i)+\alpha\phi_n(i,\alpha,k)]\geq 0$$

for $k=0, 1, 2, \ldots, s$ and that the optimal policy in period n has the control limit form given by display (1). Then

$$\sum_{j=0}^{\infty}\{P_{ij}^{m}[K_{n+1}(j)+\alpha\phi_n(j,\alpha,m)]-P_{ij}^{m'}[K_{n+1}(j)+\alpha\phi_n(j,\alpha,m')]\}$$

is nondecreasing in i and the optimal policy in period $n+l$ has the control limit form given by display (1).

COROLLARY: If $K_1, \ldots, K_n$ are sufficiently convex such that

$$\frac{\Delta^2[K_{n+1}(i)+\alpha\phi_n(i,\alpha,k)]}{\Delta i^2}\geq 0$$

for $n=0, 1, \cdots, n-1$ and all $k=0, 1, \cdots, s$, then the optimal policies for the first n periods have the control limit form.

REFERENCES

[1] Balachandran, K. R., "Control Policies for a Single Server System," Management Science, *19*, 1013–1018 (1973).

[2] Bell, C. E., "Characterization and Computation of Optimal Policies for Operating an M/G/1 Queuing System with Removable Server," Operations Research, *20*, 2080–2180 (1972).

[3] Heyman, D. P., "Optimal Operating Policies for M/G/1 Queuing Systems," Operations Research, *16*, 362–382 (1968).

[4] Magazine, M. J., "Optimal Control of Multi-channel Service Systems," Naval Research Logistics Quarterly *18*, 177–183 (1971).

[5] Sobel, M. J., "Optimal Average-Cost Policy for Queue with Start-up, and Shut-down Costs," Operations Research, *17*, 145–162 (1969).

[6] Zacks, S. and M. Yadin, "Analytic Characterization of the Optimal Control of a Queueing System," Journal of Applied Probability, 617–633 (1970).

CYCLICAL JOB SEQUENCING ON MULTIPLE SETS OF IDENTICAL MACHINES*

Helman I. Stern

Ben Gurion University of the Negev
Beersheva, Israel

Edgardo P. Rodriguez

World Bank
Washington, D.C.

Merlin L. Utter

Proctor and Gamble
Cincinnati, Ohio

ABSTRACT

The problem posed in this paper is to sequence or route n jobs, each originating at a particular location or machine, undergoing $r-1$ operations or repairs, and terminating at the location or machine from which it originated. The problem is formulated as a 0–1 integer program, with block diagonal structure, comprised of r assignment subproblems; and a joint set of constraints to insure cyclical squences. To obtain integer results the solutions to each subproblem are ranked as required and combinations thereof are implicitly enumerated. The procedure may be terminated at any step to obtain an approximate solution. Some limited computational results are presented.

INTRODUCTION

Much attention has been devoted to the classical n job-m machine shop scheduling problem. In most investigations each job is given a prespecified technological ordering. Less attention has been given to the problem of job processing options. For example, a job may be processed on machine A or machine B but not both. The problem posed in this paper is to sequence or route n jobs, each originating at a particular location or machine, undergoing $r-1$ ordered operations or repairs, in which the s^{th} operation may be performed on only one of n similiar machines; after which each job terminates at the machine from which it originated. The problem may also be visualized as the

*This research was partially supported by National Science Foundation grant No. GK–27836.

routing of a fleet of n vehicles or ships with each vehicle loading a commodity of type 1 at its origin location, delivering this commodity to any city in group 1, reloading a commodity of type 2 to be delivered to any city in group 2, etc.; such that each city in a group is visited exactly once. On the final leg of the journey each vehicle either returns empty or discharges a commodity of type r at its origin location. The problem is also a special case of the m traveling salesman problem requiring m disjoint closed tours subject to a visitation constraint where each salesman must visit one city from each group in a specified order.

To give the problem concreteness let there be n jobs and r different types of machines with n identical machines of each type

$$(s=1, 2, \ldots, r).$$

Let the i^{th} machine of type s be represented as $m_i(S)$. Each job k is to initiate and terminate its sequence on the same machine of type 1, say machine $m_k(1)$. Job k must be processed on exactly one machine of each of the machine types 2, 3, . . ., r, in increasing order. No machine may process more than one job. Thus, the k^{th} job has the technological ordering

$$\{[m_k(1)], [m_1(2) \text{ or } m_2(2) \ldots \text{or } m_n(2)], \ldots, [m_1(s) \text{ or } m_2(s) \ldots \text{or } m_n(s)], \ldots,$$
$$[m_1(r) \text{ or } m_2(r) \ldots \text{or } m_n(r)], [m_k(1)]\}.$$

A feasible solution to this problem must have each machine assigned to one and only one job. An example of a feasible solution for 3 jobs and 4 machine types is shown in Figure 1, and consists of 3 disjoint cycles each comprised of four arcs.

For this particular problem there are $(3!)^3=216$ feasible solutions. In general for an n job—r machine type problem there are $(n!)^{r-1}$ feasible solutions. If the cost of transfering a job of any kind from machine $m_i(s)$ to $m_j(s+1)$ is given as $c_{ij}(s)$ the problem becomes one of finding an efficient algorithm that searches the set of feasible solutions and selects that solution (or solutions) that minimizes the total cost of sequencing all jobs. It is assumed that the processing costs on any machine of a given type are job independent. Thus, if a processing cost of $c_i(s)$ represents the cost to

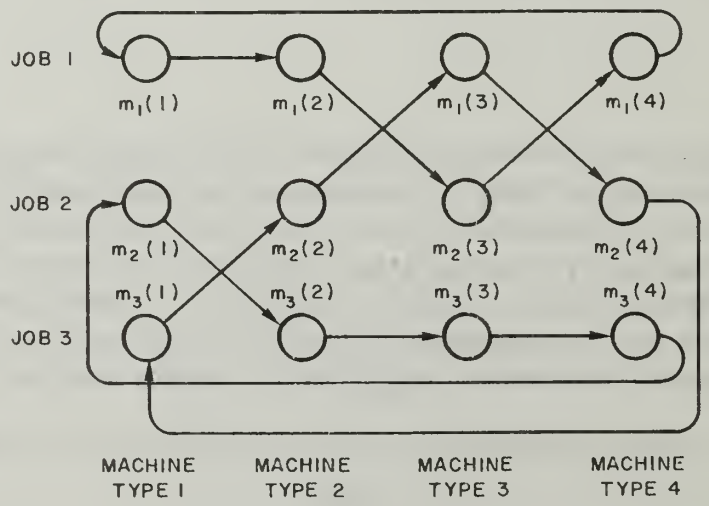


FIGURE 1.—Cyclical sequencing of three jobs through four sets of machines.

process any job on machine $m_i(s)$, it may be added to the transfer cost and included in $c_{ij}(s)$ without loss of generality. The problem is formulated as a 0–1 integer program, with block diagonal structure comprised of r assignment subproblems, and a joint set of constraints that insure closed tours. Since all variables are 0–1, the traditional Dantzig-Wolfe decomposition scheme is precluded. To insure integer results the solutions to each assignment subproblem are ranked as required and combinations thereof are implicitly enumerated in a branch and bound scheme. The procedure may be terminated at any step to obtain a feasible solution along with bounds on its accuracy. Some limited computational results are given in the last section.

Related work has appeared in the literature starting with the truck dispatching problem of Dantzig and Ramser [1], Clark and Wright [2], and Newton and Thomas [3], in which all n tours originate and terminate at a single city with the number of cities visited by each tour implicitly conditional on the order of visits (due to demand variations at each city). Krolak, Felts, and Nelson [4] report on the work of Newton and Thomas in which multi-origin, single destination routes are scheduled. Bellmore, Liebman, and Marks [5] consider a multi-origin problem in which each tour is open but with ordered node set visitations. Svestka and Huckfield [6] solve an n closed tour problem through m nodes such that n sorties of any length commence and terminate at a single node. Srivastava, Kumar, Garg, and Sen [7] consider a single closed tour visiting r cities, one from each of r ordered sets of cities, such that each set has a cardinality of at least one.

MATHEMATICAL FORMULATION

The problem may be formulated as a 0–1 minimum cost network flow problem with closed tour side conditions. The underlying network is depicted in Figure 2 where the node set N_s represents the set of n machines of type s. The node sets are ordered from left to right with N_1 repeated and designated as N_{r+1} as a visual aid. The only arcs in the network are those emanating from each set of nodes to the next higher set of nodes. The arcs from N_r to N_{r+1} are return arcs required to complete cycles. Let the machine sequence of each job start from its origin node in N_1, proceed through a single node from each of the remaining ordered sets, N_2, N_3, . . ., N_r, and end at its origin node repeated in N_{r+1}. In Figure 2 the only arcs shown are those possible for a cyclical

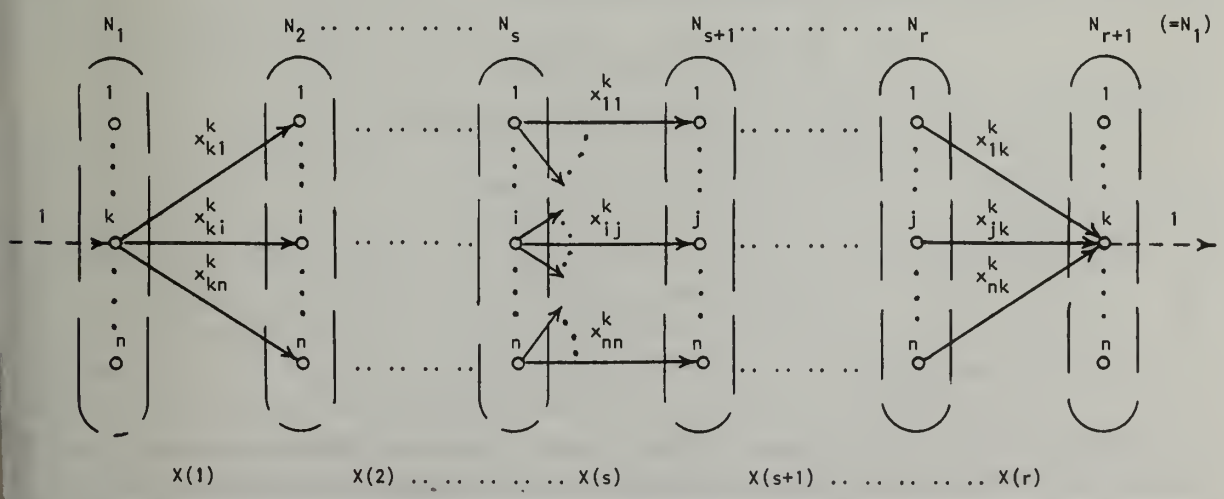


FIGURE 2.—Network representation of problem (only flows for job k are shown).

machine sequence for job k. The variables, $x_{ij}^k(s)$, represent the amount of job k flowing from node i in N_s to node j in N_{s+1}. The set of all flows from N_s to N_{s+1} is represented as $X(s)$. To insure full job assignments to machines let $x_{ij}^k(s)$ be restricted to a unit value if the k^{th} job is sequenced on machine j in N_{s+1} directly after machine i in N_s. Otherwise, the flow will be restricted to zero. Total flows through each node in th network will be restricted to unity so that exactly one job will be sequenced through each machine. In addition, only the k^{th} job may flow through node k of N_1. Thus, one may drop the superscripts from $X(1)$ and $X(r)$ and reduce the number of flow variables to those described by y_{ki} (1) and y_{ik} (r). (This is an arbitrary stipulation, as any one to one correspondence between the n jobs and n machines or origins in N_1 will suffice.) This final restriction insures that the k^{th} job originates and terminates its sequence at node k of sets N_1 and N_{r+1}. In addition, define 0–1 variables $y_{ij}(s)$ for $s=2, 3, \ldots, r-1$ which represent the joint flows from each arc (i, j) in (N_s, N_{s+1}). Let $c_{ij}(s)$ represent the cost (distance) of any job transferred from i in N_s to j in N_{s+1}, and Z represent the total cost incurred by all job transfers. Then the problem of minimizing the total job sequencing cost may be formulated as a minimum cost flow problem with 0–1 integer flows and closed tour side constraints. The mathematical formulation of this problem is shown below as problem (P).

Integer Program (P)

$$\text{Min } Z=\sum_{(k,\,i)\epsilon(N_1 N_2)} c_{ki}(1)y_{ki}(1)+\sum_{s=2}^{r-1}\sum_{(i,\,j)\epsilon(N_s,\,N_{s+1})} c_{ij}(s)y_{ij}(s)+\sum_{(j,\,k)\epsilon(N_r,\,N_{r+1})} c_{jk}(r)y_{jk}(r)$$

Subject to:

$$\left.\begin{aligned}(1)\quad &\sum_{j\epsilon N_{s+1}} y_{ij}(s)=1,\ i\epsilon N_s\\ (2)\quad &\sum_{i\epsilon N_s} y_{ij}(s)=1,\ j\epsilon N_{s+1}\end{aligned}\right\} s=1, \ldots, r$$

$$(3)\qquad y_{ki}(1)-\sum_{j\epsilon N_3} x_{ij}^k(2)=0,\ (k,\ i)\epsilon(N_1,\ N_2)$$

$$(4)\qquad y_{ij}(2)-\sum_{k=1}^{n} x_{ij}^k(2)=0,\ (i,\ j)\epsilon(N_2,\ N_3)$$

$$\left.\begin{aligned}(5)\quad &\sum_{i\epsilon N_{s-1}} x_{ij}^k(s-1)-\sum_{i\epsilon N_{s+1}} x_{ji}^k(s)=0,\ j\epsilon N_s\\ (6)\quad &y_{ij}(s)-\sum_{k=1}^{n} x_{ij}^k(s)=0,\ (i,\ j)\epsilon(N_s,\ N_{s+1})\end{aligned}\right\} s=3, \ldots, r-1$$

$$(7)\qquad y_{jk}(r)-\sum_{i\epsilon N_{r-1}} x_{ij}^k(r-1)=0,\ (j,\ k)\epsilon(N_r,\ N_{r+1})$$

$$(8)\qquad y_{ij}(s)=0,\ 1;\ (i,j)\epsilon(N_s,\ N_{s+1})\qquad s=1, \ldots, r$$

$$(9)\qquad x_{ij}^k(s)=0,\ 1;\ (i,\ j)\epsilon(N_s,\ N_{s+1})\qquad s=2, \ldots, r-1\qquad k=1, \ldots, n$$

Equations (1), (2), and (8) represent r independent assignment problem constraints insuring that each node has unit flow. Together they provide n disjoint job sequences from nodes in N_1 to nodes in N_{r+1} not necessarily closed. The interstage coupling constraints (3), (4), (5), (6), (7), and

(9) provide the conditions necessary for all tours to be closed (cyclical sequences). Upon rearrangement, the entire set of constraints can be shown to exhibit a block diagonal structure.

DECOMPOSITION OF (P)

Let (B) represent the set of closed tour coupling constraints. Removal of (B) from (P) allows the remaining portion of (P) to be decoupled into r independent assignment problems. Let (A_s) represent the s^{th} stage assignment problem from machines in N_s to machines in N_{s+1}. Thus, (A_s) is of the form,

Minimize

$$Z_s = \sum_{(i,j)\epsilon(N_s, N_{s+1})} c_{ij}(s) y_{ij}(s)$$

Subject to:

$$\sum_{j\epsilon N_{s+1}} y_{ij}(s) = 1, \; i\epsilon N_s$$

$$\sum_{i\epsilon N_s} y_{ij}(s) = 1, \; j\epsilon N_{s+1}$$

$$y_{ij}(s) = 0, 1; \; (i, j)\epsilon(N_s, N_{s+1})$$

Let (A) represent the union of all r assignment problems with a feasible solution defined by $[Z, Y]$ such that;

$$Z = \sum_{s=1}^{r} Z_s, \; Y = [Y_1, \ldots, Y_s, \ldots, Y_r],$$

where Y_s is a feasible solution to (A_s) and Z_s is its associated cost.

Note that an optimal solution $[Z^0, Y^0]$ to (A) provides a lower bound to (P). If, in addition, there exists a solution $[X, Y^0]$ to (B), then $[Z^0, Y^0, X]$ is also optimal for (P). Otherwise, one may enumerate all feasible solutions to (A) and select, among those that also provide a feasible solution to (B), the one with the lowest objective value. Such a process, however, requires the computation of $(n!)^r$ solutions to (A) and an equal number of feasibility tests in (B).† A problem with $n=7$ and $r=3$ requires the determination 6.4×10^{14} feasible solutions to (A). Hence, an implicit enumeration branch and bound procedure is proposed. This procedure requires an efficient technique for determining if proposed solutions to (A) satisfy the side conditions in (B), as well as a method of constructing upper bounds on (P).

PROPOSITION 1: Testing Y for Feasible Closed Tours

If (Z^*, Y^*) is a feasible solution to (A), then (B) has either a unique solution $[X^*, Y^*]$ and Z^* provides an upper bound for (P), or there does not exist a solution to (B) corresponding to Y^*.

This result follows from the staircase structure of the constraint set (B) where (2) and (4) link stages 1 and 2, (5) and (6) link stages s and $s+1$, for $s=2, \cdots, r-1$, and (7) links stages $r-1$ and r. Since each assignment solution provides exactly n variables with unit values and $n(n-1)$ variables at zero, the $Y_1^*, \cdots Y_{r-1}$ determine uniquely $X_1^*, \cdots X_{r-1}$. If Y_r^*, X_{r-1} also satisfy (7), Z^* provides an upper bound for (P). Otherswie, there does not exist a solution to (P) corresponding to Y^*. Intuitively, one is projecting the paths of all n jobs, stagewise, from their origin nodes in N_1 to their terminating nodes in N_{r+1}. Constraint (7) provides the test to determine if the k^{th} job originates and terminates at the k^{th} node in N_1. In lieu of sequentially solving the set of coupling equations for the $x_{ij}^k(s)$ one may incorporate a composite function on the subscripts

†Note that all feasible solutions to the assignment problem are basic solutions.

of the Y^* solution for an easily computerized closed tour test. Moreover, if Y^* does not provide a closed tour solution, fixing Y^* for any $r-1$ stages yields a forced feasible solution for the remaining r^{th} stage providing a feasible upper bound on (P). This observation yields the following proposition:

PROPOSITION 2: Construction of Revised Feasible Solution

Given (Z^*, Y^*), a solution to (A), which does not provide a feasible solution for (B). One may obtain a feasible solution to (B) to by constructing a revised solution

$$Y^{*s}=[Y_1^*, \ldots, \hat{Y}_s^*, \ldots Y_r^*]$$

where $\hat{Y}^*_s$ is the unique s^{th} stage solution to (A_s) which allows the construction of a feasible solution to (B) (a revised solution) and provides a feasible upper bound of Z^{*s} on (P). By computing a set of r revised upper bounds Z^{*s}, $s=1, \ldots, r$, for (P) the tightest upper bound, $\bar{Z}^*$, may be determined as,

$$\bar{Z}^*=\underset{s=1,\ldots,r}{\text{Min}}\ Z^{*s}$$

Example of Revised Solutions

Consider a solution shown pictorially in Figure 3a for a three machine problem with three jobs. This solution is feasible for (A) but infeasible for (B). The three uniquely determined revised solutions (feasible for (B)) are shown in Figures 3b, 3c and 3d.

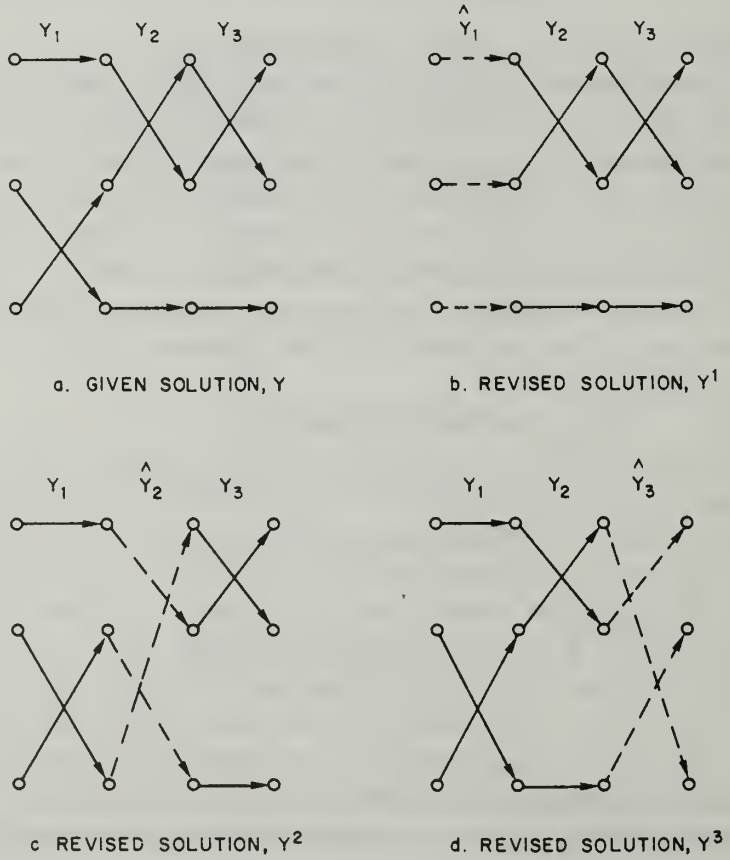


FIGURE 3.—Example of revised solutions.

TERMINOLOGY

The results of propositions 1 and 2 provide the constructive basis for determining solutions to (A) and upper bounds on (P). To facilitate the desire to implicitly enumerate all potential solutions of (A) the following definitions are introduced:

1. The Ranked Assignment, Y_{st}:

The assignments to (A_s) may be ordered by non-decreasing values of Z_s, such that Y_{st} represents the t^{th} ranked assignment and Z_{st} its associated cost. The ranking index $t=1, \ldots, n!$ is selected by the rule,

$$\forall a, b \epsilon \{1, \ldots, n!\} \ni a > b, Z_{sa} \geqq Z_{sb}$$

2. A Node, Y^k:

Any feasible solution to (A) say, the one identified as the k^{th} solution, may be represented as a node, Y^k, where,

$$Y^k = \{Y^k_{u_1}, \ldots, Y^k_{st_s}, \ldots, Y^k_{rt_r}\}$$

$$\ni t_s \epsilon \{1, \ldots, n!\}, s=1, \ldots, r.$$

Where clear, a node may be represented as an ordered set of assignments, each assignment represented by the integer equal to its rank, i.e.,

$$(t_1, \ldots, t_s, \ldots t_r)$$

3. Set of Solutions:

Let $P_{u_1}, \ldots, u_s, \ldots u_r$ be the set of all solutions to (A) less the (u_s-1) best solutions for (A_s), $s=1, \ldots, r$, i.e.,

$$P_{u_1, \ldots, u_s, \ldots u_r} = \{(t_1, \ldots, t_s, \ldots, t_r)/t_s \geqq u_s, s=1, \ldots r\}.$$

Thus, $P_{1, \ldots, 1, \ldots, 1}$ is the set of all feasible solutions to (A). (When clear we shall refer to the set of all feasible solutions as P.)

4. Cost of Node Y^k, Z^k:

The sum of the costs of each of the assignments in Y^k is

$$Z^k = \sum_{s=1, \ldots, r} Z_{st_s}{}^k$$

5. A Feasible Node, Y^k:

A node, Y^k, shall henceforth be called feasible if and only if it provides a feasible solution to (P). (This may be determined by the results of Proposition 1.)

6. A Revised Node, Y^{ks}:

A revised node, Y^{ks}, is the node obtained from an infeasible node Y^k by replacing the s^{th} assignment $Y^k_{st_s}$ by the unique assignment $\hat{Y}^k_{su_s}$ such that the new node becomes feasible: $(u_s \neq t_s)$

7. Cost of Revised Node Y^{ks}, Z^{ks}:

The cost associated with the revised node Y^{ks} is

$$Z^{ks} = Z^k - Z^k_{st_s} + \hat{Z}^k_{su_s}.$$

8. Upper Bound Associated With Node Y^k, $\bar{Z}^k$:

An upper bound on (P) associated with node Y^k may be determined as Z^k if Y^k is feasible, or from its set of revised nodes. (See Proposition 2.)

If Y^k is infeasible

$$\bar{Z}^k = \min_{s=1,\ldots,r} Z^{k_s}$$

i.e.,

$$\bar{Z}^k = \begin{cases} Z^k, \text{ if } Y^k \text{ is feasible} \\ \min_{s=1,\ldots,r} Z^{ks}, \text{ if } Y^k \text{ is infeasible} \end{cases}$$

9. Dominance Between Nodes:

The node $(u_1, \ldots, u_s, \ldots u_r)$ is said to be dominated by the node $(v_1, \ldots, v_s, \ldots v_r)$ if

$$u_s \geqq v_s, \; s = 1, \ldots r.$$

It follows from 1 and 4 above that if node Y^k is dominated by node Y^l then

$$Z^k \geqq Z^l.$$

In Figure 4 all nodes below the dotted line are dominated by the node (1, 2, 1).

IMPLICIT ENUMERATION OF SOLUTIONS THROUGH BRANCH AND BOUND OF SETS OF RANKED ASSIGNMENTS

In this section the general rational for searching the solution set P is presented in terms of a branch and bound algorithm. The algorithm is initiated by generating the node Y^1 (1, . . ., 1, . . ., 1) and testing it for feasibility using the results of Proposition 1†. If Y^1 is feasible then it represents an optimal solution to (P) with optimal cost Z^1. Otherwise, Z^1 offers a least lower bound to (P), in which case an upper bound may be constructed from Y^1 by employing the result of Proposition 2. The scheme for generation of future nodes shall incorporate the assignment ranking procedure of K.G. Murty [8].

It is useful to present the scheme in terms of the tree diagram shown in Figure 4 (for a three stage problem) with each node represented as a point and arrows between points representing branching from a predecessor node to a successor node.

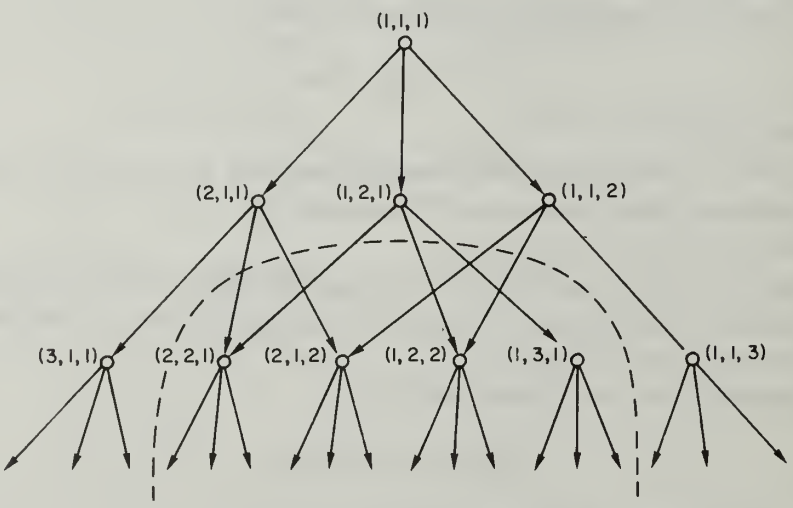


FIGURE 4.—Illustrative tree for a three stage problem.

†According to Definition 2 this node represents the first ranked solutions of all assignment subproblems.

Branches in the tree are limited to those from a node of the form $(t_1, \ldots, t_s, \ldots t_r)$ to successor nodes of the form

$$(t_1+1, \ldots, t_s, \ldots t_r), \ldots, (t_1, \ldots, t_s+1, \ldots, t_r), \ldots, (t_1, \ldots, t_s, \ldots t_r+1).$$

Any of the successor nodes, say the s^{th} one, may easily be determined from its predecessor through the construction of the next best ranked assignment Y_{s,t_s+1} for s^{th} stage assignment problem. To reduce the information storage requirements the ranked assignments are exhausted one stage at a time (to be clarified subsequently).

Associated with each node in the tree, say the k^{th}, are its cost Z^k and an upper bound $\overline{Z}^k$. If the node is feasible then Z^k provides an upper bound and is set equal to $\overline{Z}^k$, otherwise an associated upper bound $\overline{Z}^k$ is determined through construction of revised nodes (See Definition 8).

The best solution determined for a tree with k nodes may be defined as the least upper bound

$$Z_u^k = \underset{l=1,\ldots,k}{\text{Min.}} \overline{Z}^l$$

In the $k+1$ *st* step the best upper bound may be updated by the rule,

$$Z_u^{k+1} = \begin{cases} Z_u^k, \text{ if } \overline{Z}^{k+1} \geqq Z_u^k \\ \overline{Z}^{k+1}, \text{ if } \overline{Z}^{k+1} < Z_u^k \end{cases}$$

To check for the existence of dominated nodes at step $k+1$ (the dominance test) compare the cost, Z^{k+1}, associated with the node Y^{k+1} (feasible or infeasible), with Z_u^{k+1}. If $Z^{k+1} \geqq Z_u^{k+1}$ the set of nodes dominated by Y^{k+1} may be removed from the tree as they will exhibit costs not less than those of the best solution found thus far. As more nodes are examined, the tree is reduced in size until a point is reached when all nodes in the reduced tree have been examined. If this point is reached at step q, then Z_u^q is the optimal cost. This is true since Z_u^q represents the least upper bound on all nodes explicitly examined, and is a better solution than those nodes implicitly examined through placement in the dominated set. The algorithm will terminate in a finite number of steps since the number of nodes in the initial tree is $(n!)$, and unexamined nodes in the tree are evaluated at each step with at least one being removed. This may be shown in set theoretic notation as follows:

Let T^k = the set of nodes in the tree at step k (reduced tree).

E^k = set of nodes in T^k examined in previous steps.

$\overline{E}^k$ = set of nodes not yet examined in T^k.

D^k = set of nodes dominated by nodes in E^k.

Thus, at step k of the algorithm the set of nodes P may be partitioned into three nonintersecting sets, i.e.,

$$P = T^k \cup D^k = E^k \cup \overline{E}^k \cup D^k$$

On the $k+1$ *st* iteration a node Y^{k+1} is selected from $\overline{E}^k$ and placed in E^k. If Y^{k+1} determines a set of dominated nodes those dominated nodes not in E^k, say $D(Y^{k+1})$, will also be deleted from $\overline{E}^k$ and added to D^k.

Thus,

$$|E^{k+1}|=|E^k|+1$$

$$|\bar{E}^{k+1}|=|\bar{E}^k|-1-|D(Y^{k+1})|$$

$$|D^{k+1}|=|D^k|+|D(Y^{k+1})|$$

and

$|\bar{E}^{k+1}|<|\bar{E}^k|$. Since $|E^1|=P$ and P is finite, termination occurs at some point $k=q \ni \bar{E}^q=\emptyset$.

A sketch of the algorithm follows

THE ALGORITHM

Initialization:

Let $k=1$, $\bar{E}^k=P$, $E^k=D^k=\emptyset$, $Z_u{}^k=\infty$

Step 1: Select a node Y^k from $\bar{E}^k$. If $\bar{E}^k=\emptyset$,

STOP: OPTIMAL SOLUTION IS $Z_u{}^k$.

Step 2: Compute Cost of Y^k

$$Z^k=\sum_{s=1,\ldots r} Z^k_{st_s}$$

Step 3: Feasibility check

$$\text{Let } \gamma^k=\begin{cases}1, \text{ if } Y^k \text{ is feasible}\\ 0, \text{ if } Y^k \text{ is infeasible}\end{cases}$$

Step 4: Upper bound associated with Y^k

$$\bar{Z}^k=\begin{cases}Z^k, \text{ if } \gamma^k=1\\ \underset{s=1,\ldots r}{\text{Min}}\ Z^{ks}, \text{ if } \gamma^k=0\end{cases}$$

Step 5: Update best solution

$$Z_u{}^k=\begin{cases}Z_u{}^k, \text{ if } \bar{Z}^k \geqq Z_u{}^k\\ \bar{Z}^k, \text{ if } \bar{Z}^k < Z_u{}^k\end{cases}$$

Step 6: Determine dominated nodes at Y^k

$$D(Y^k)=\begin{cases}P_{t_1,\ldots t_s \ldots t_r}, \text{ if } \gamma^k=1 \text{ or } \bar{Z}^k \geqq Z_u{}^k\\ \emptyset, \text{ if } \gamma^k=0\end{cases}$$

where $P_{t_1,\ldots t_s \ldots t_r \ldots}$, is the set of nodes dominated by the node $Y^k=\{Y^k_{1t_1} \ldots Y^k_{st_s} \ldots Y^k_{rt_r}\}$

Step 7: Update Solution Sets

$$E^{k+1}=E^k+Y^k$$

$$E^{k+1}=E^k-Y^k-D(Y^k)$$

$$D^{k+1}=D^k+D(Y^k)$$

Return to Step 1

The selection of a node from $\bar{E}^k$ deserves further explanation. In the actual programming of the algorithm (see [9]) the set operations are performed implicitly, and hence no bookkeeping

requirements are necessary to record the set element transfer operations in Step 7. Moreover, the elements in $\bar{E}^k$ are generated as selected. This selection decision is designed to fulfill two objectives, (i) reduce storage in the assignment ranking routine, and (ii) accelerate the generation of dominated nodes through the exploitation of all computed assignments thus far.

The selection scheme may be described with the aid of the list of ranked assignments shown in Table 1.

TABLE 1.—*List of Ranked Assignments*

Ranked assignment	Subproblem				
	A_1	· · ·	A_s	· · ·	A_r
1	Y_{11}	· · ·	Y_{s1}	· · ·	Y_{r1}
2	Y_{12}	· · ·	Y_{s2}	· · ·	Y_{r2}
·	·		·		·
·	·		·		·
·	·		·		·
t	Y_{1t_1}	· · ·	Y_{st_s}	· · ·	Y_{rt_r}
·	·		·		·
·	·		·		·
·	·		·		·
$n!$	$Y_{1n!}$	· · ·	$Y_{sn!}$	· · ·	$Y_{rn!}$

A node Y^k is an ordered set of elements selected from Table 1 in such a manner that one element is selected from each column in the list and ordered column wise. Hence, the list represents the basic data for generating all $(n!)^r$ solutions in P. Any column in the list may be generated using Murty's assignment ranking technique. This technique requires a considerable amount of stored data from the $t-1$ first assignments to determine the t^{th} ranked assignment. Thus, it is efficient to exhaust the generation of assignments one column at a time.

One proceeds from the origin node $(1, \cdots, 1) \cdots 1)$ and generates successive nodes by ranking the assignments of the first stage assignment problem until it is determined that a node using the next assignment in the first column of the list is feasible or provides inferior solutions by the law of dominance. (See Step 6.) This is illustrated in Figure 5 as path 1 for a three stage example. One then backtracks and starts to rank the assignments of the second stage. The third stage assignments are not started until all undominated nodes involving combinations of already ranked solutions in stage 1 are examined. These combinations for the example illustrated in Figure 4 are shown in paths 2 and 3. One continues in this fashion until all stages have been ranked. The details of exhausting all combinations of previously generated assignments may be found in reference [9] which is available from the authors upon request. The efficiency of this scheme is demonstrated in the following three stage example where only 9 out of a possible 216 nodes are explicitly examined. Moreover, the scheme may be terminated at any step to obtain an approximate feasible solution and a bound on the cost error of this solution through comparison with the lower bound found in the first step.

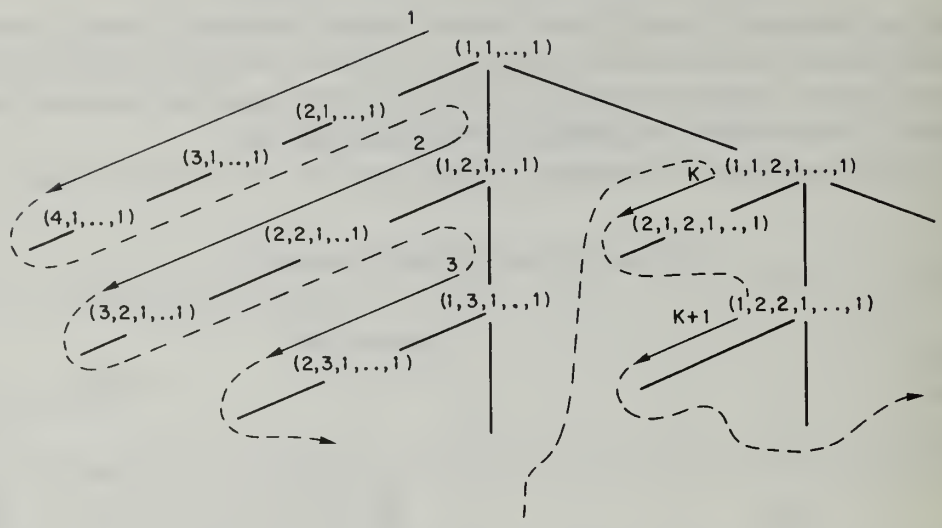


FIGURE 5.—A tree to illustrate the sequence of paths followed during the branching scheme.

EXAMPLE

The cost matrices for this three job-three machine stage example and the associated list of ranked assignments are shown in Tables 2 and 3, respectively.

TABLE 2.—*Cost Matrices for Three Stage Example*

Stage 1			Stage 2			Stage 3		
4	6	2	3	5	9	4	8	0
1	8	10	10	0	1	2	0	2
5	2	5	2	8	5	7	0	0

TABLE 3.—*Ranked Assignments for Three Stage Example*

Stage 1			Stage 2			Stage 3		
Rank	Assignment	Cost	Rank	Assignment	Cost	Rank	Assignment	Cost
1	312	5	1	231	8	1	312	2
2	213	12	2	123	8	2	123	4
3	321	15	3	321	11	3	132	6
4	132	16	4	132	12	4	321	7
5	123	17	5	213	20	5	213	10
6	231	21	6	312	27	6	231	17

The assignments in Table 3 are represented as a permutation of the sequence (1, 2, 3). The computational and branching results for each step of the algorithm are shown in Table 4 and Figure 6, respectively.

TABLE 4.—*Nodal Information at Each Step of the Algorithm for Three Stage Example*

Step	Node	Cost	Upper bound	Min upper bound	Feasibility	Set of dominated nodes
k	Y^k	Z^k	$\overline{Z}^k$	$Z_u{}^k$	γ^k	$D(Y^k)$
1	(1, 1, 1)	15	23	23	0	$\emptyset$
2	(2, 1, 1)	22	27	23	0	$\emptyset$
3	(3, 1, 1)	25	25	23	0	$P_{3,\,1,\,1}$
4	(1, 2, 1)	15	31	23	0	$\emptyset$
5	(2, 2, 1)	24	25	23	0	$P_{2,\,2,\,1}$
6	(1, 3, 1)	18	22	22	0	$\emptyset$
7	(1, 4, 1)	19	23	22	0	$\emptyset$
8	(1, 5, 1)	27	22	22	0	$P_{1,\,5,\,1}$
9	(1, 1, 2)	17	17	17	1	$P_{1,\,1,\,2}$

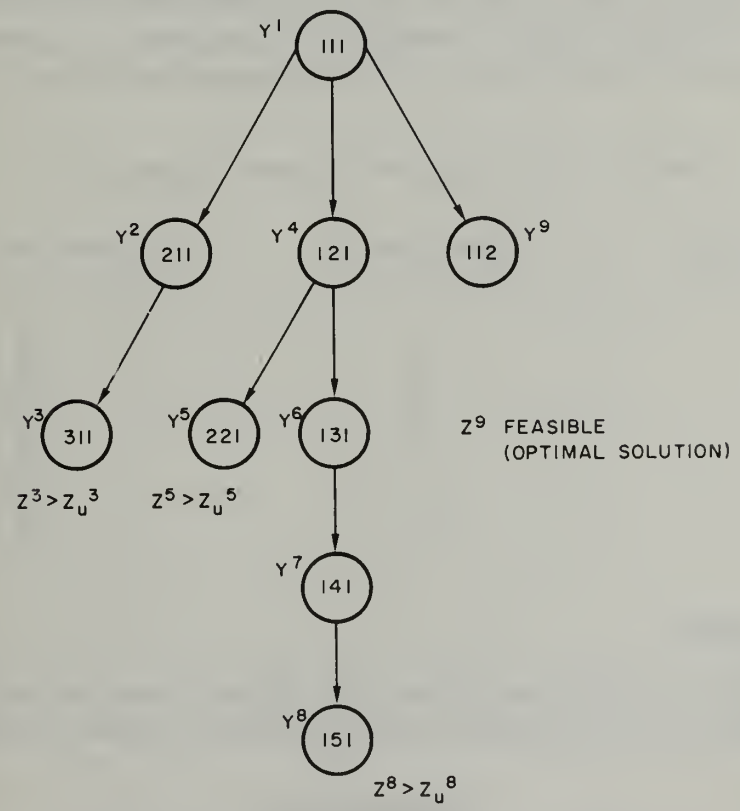


FIGURE 6.—Tree illustration of three stage example.

The optimal cyclical job sequences are shown in Figure 7.

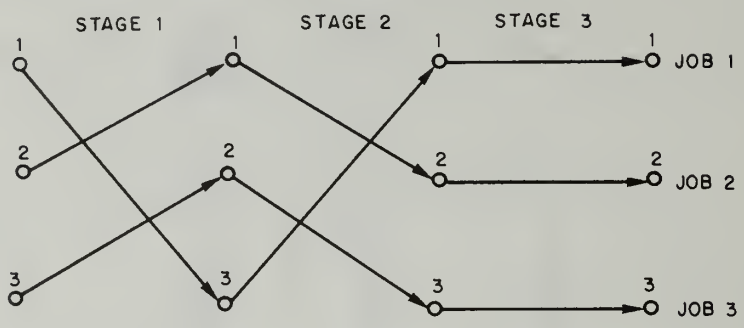


FIGURE 7.—Optimal cyclical job sequences for three stage example.

It is noted that of the 6 ranked assignments for each stage only the first 3, 5 and 2 ranked assignments for stages 1, 2 and 3, respectively had to be computed in the example.

A computer program has been written in Fortran IV for the IBM 360/50 to apply the algorithm for r equal three [9]. Eight sample problems were solved using this code with execution times shown in Table 5. The reader should be cautioned that these times are expected to increase as additional stages are added beyond three, although no computational evidence has been accumulated as yet.

TABLE 5.—*Computer Results for Sample Problems* ($r=3$)

Problem number	Problem size (n)	Execution time (seconds)
1	3	8.5
2	4	8.8
3	4	9.5
4	5	8.7
5	5	14.5
6	7	12.3
7	7	21.2
8	10	36.1

ACKNOWLEDGMENTS

The authors wish to give their appreciation to Dr. Richard Francis of Ohio State University for securing permission to use the assignment ranking computer program referenced in [10].

REFERENCES

[1] Dantzig, G. B. and J. H. Ramser, "The Truck Dispatching Problem," Management Science *6*, 80–91 (1959).

[2] Clarke, G. and J. W. Wright, "Scheduling of Vehicles from a Central Depot to a Number of Delivery Points," Operations Research, *12*, 568–81 (1964).

[3] Newton, R. M. and W. H. Thomas, "Design of School Bus Routes by Computer," Socio-Economic Planning Sciences, *3*, 75–85 (1969).

[4] Krolak, P., W. Felts and J. Nelson, "A Man-Machine Approach Toward Solving the Generalized Truck-Dispatching Problem," Transportation Science, *6*, 149–170 (1972).

[5] Bellmore, M., J. C. Liebman and D. H. Marks, "An Extension of the (Szwarc) Truck Assignment Problem," Naval Research Logistics Quarterly, *19*, 91–99 (1972).

[6] Svestka, J. A. and V. E. Huckfeldt, "Computational Experience with an M-Salesman Traveling Salesman Algorithm" Management Science, *7*, 790–799 (1973).

[7] Srivastava, S. S., S. Kumar, R. C. Garg and P. Sen, "Generalized Traveling Salesman Problem Through n Sets of Nodes," Canadian Operational Research Society, *7*, 97–101 (1969).

[8] Murty, K. G., "An Algorithm for Ranking All the Assignments in Order of Increasing Costs," Operations Research, *16*, 682–687 (1968).

[9] Stern, H. I., P. Rodriguez and M. L. Utter, "The M-Traveling Salesman Problem with Ordered Visits," Operations Research and Statistics Paper No. 37–71–P5, Rensselaer Polytechnic Institute, Troy, New York (1971).

[10] Fluharty, E., "Solving the Quadratic Assignment Problem by Ranking Assignments," Masters Thesis, Ohio State University (1969).

JOHNSON'S APPROXIMATE METHOD FOR THE 3 × n JOB SHOP PROBLEM

Wlodzimierz Szwarc and George K. Hutchinson

School of Business Administration
University of Wisconsin-Milwaukee
Milwaukee, Wisconsin

ABSTRACT

The effectiveness of Johnson's Approximate Method (JAM) for the 3 × n job shop scheduling problems was examined on 1,500 test cases with n ranging from 6 to 50 and with the processing times A_i, B_i, C_i (for item i on machines A, B, C) being uniformly and normally distributed. JAM proved to be quite effective for the case $B_i \leq \max(A_i, C_i)$ and optimal for $B_i \leq \min(A_i, C_i)$.

1. INTRODUCTION

The 3 × n job shop problem can be defined as follows. Machines A, B, C process each of the n items $1, 2, \ldots, n$ in the order $A\,B\,C$. Given the processing times (A_i, B_i, C_i for item i on machines A, B, C respectively), the problem is to find a sequence minimizing the timespan. S. M. Johnson [2] provided a very quick method for solving the 2 × n problems which, when adopted to the 3 × n case, reads as follows.

Solve instead a 2 × n problem assuming the processing times for item i on the first and second machines to be A_i+B_i and B_i+C_i respectively. The resulting sequence p is not necessarily optimal, however.

R. J. Giglio and H. M. Wagner [1] tested Johnson's Approximate Method (JAM) on twenty 3 × 6 trials (with processing times uniformly distributed between 1 and 30). In 9 out of 20 trials the optimal sequence was produced. To check whether p is optimal they had to enumerate *all n!* permutations (in general there is no other way) and find the respective elapsed time.

Let $T(p)$ and $t(p)$ be the timespans for sequence p in the 3 × n and 2 × n ($A+B$, $B+C$) problems respectively. W. Szwarc established in [3] sufficient optimality conditions for JAM. Sequence p is optimal for the 3 × n problem if

$$T(p)=t(p)-\sum_{i=1}^{n} B_i \tag{1}$$

REMARK: The right hand side of (1) is actually the lower bound of the minimal timespan of the three machine problem.

The efficiency of Johnson's two machine method makes it possible to test JAM (via (1)) for practically any number of items. There were examined 1,500 $3 \times n$ trials with n ranging from 6 to 50. The processing times were being drawn from a uniform and a normal distribution. The results show that JAM is quite effective if $B_i \leq \max(A_i, C_i)$ for all i (see Tables 1 and 2). JAM always produces an optimal solution for the case $B_i \leq \min(A_i, C_i)$ which lead us to believe that a theorem confirming optimality of JAM (for this case) was true. This was subsequently proven in [4]. We feel that JAM is actually much more effective than the results indicate. This is due to the fact that (1) is too strong a condition. Moreover (see Section 3) its effectiveness seems not to depend on the number of items.

2. COMPUTATIONAL RESULTS

The program was written in ECSL and run on the UNIVAC 1110. There were performed 30 separate runs, each providing three separate streams of random numbers A_i, B_i, C_i for the 50 $3 \times n$ trials as well as calculating $T(p)$ and its lower bound.

The results are summarized by Table 1.

TABLE 1

Distribution of A_i, B_i, C_i	Restrictions on B_i	Number of items n				Row
		6	10	20	50	
Normal $\mu=50$, $\sigma=5$	$B_i \leq \max(A_i, C_i)$	43 48	46 47	47 43	—	1
Uniform 1–99	$B_i \leq \max(A_i, C_i)$	42 32	46 47	43 44	49 44	2
	None	18 14	21 16	24 25	21 24	3
	$B_i \leq \min(A_i, C_i)$	50 50	50 50	50 50	50 50	4

REMARKS: 1. Each cell registers the results of two independent 50 trial runs.

2. The numbers in the cells indicate how many times per run (i.e., out of 50) was $T(p)$ equal to its lower bound.

3. The three streams generate numbers in A_i, B_i, C_i order for each $i=1, \ldots, n$. However, in the case when B_i are restricted, the values of B_i are interchanged whenever necessary with A_i or (and) C_i to comply with the appropriate restriction.

4. The run times in CPU seconds per run (50 trials) for $n=6$, 10, 20, and 50 are 25.69, 47.10, 129.13, and 588.93 respectively.

3. STATISTICAL VERIFICATION OF THE RESULTS

The first analysis was to examine whether the random numbers generated in each run affect the results within each cell. According to the Standard Chi-Square Test based on the data in the first two rows (excluding $n=50$) the two results in a cell are not statistically significantly different on the 20% level. Moreover, there are no statistically significant differences (at the 20% level) between 1) the distributions, 2) the number of items n, 3) the interactions of n and the distributions.

Applying the same technique to the data from the last three rows, we found that 1) there are no statistically significant differences between the interactions of n and the restrictions on B_i (level 20%); 2) as expected there is a strong statistically significant difference between the various restrictions on B_i (level 0.1%).

The conclusion that the effectiveness of JAM does not depend on n seems rather surprising. Hence we decided to verify it on another test.

The ratio of the actual $T(p)$ to its lower bound (in percents) was found for each of the 800 trials from rows 2 and 3. The entries of Tables 2 and 3 register the number of trials belonging to a specific ratio and number of items category. The numbers on the left hand side in Tables 2 and 3 symbolize unit intervals. For instance 104 means a half closed interval [104, 105].

TABLE 2.—Case $B_i \leq \max(A_i, C_i)$

Ratio $\left(\frac{\text{actual time}}{\text{lower bound}}\right) \cdot 100$	Number of items				Total
	6	10	20	50	
100	83	97	97	100	377
101	7	1	3		11
102	3	1			4
103	1	1			2
104	3				3
105	0				0
106	1				1
107	0				0
108	1				1
109	0				0
110	0				0
111	0				0
112	0				0
113	0				0
114	0				0
115	1				1

TABLE 3.—No restriction on B_i.

Ratio $\left(\frac{\text{actual time}}{\text{lower bound}}\right) \cdot 100$	Number of items				Total
	6	10	20	50	
100	35	41	52	52	180
101	5	3	3	4	15
102	4	5	4	1	14
103	7	3	4	1	15
104	5	9	1	6	21
105	3	2	5	4	14
106	4	6	2	4	16
107	2	2	3	2	9
108	1	3	1	2	7
109	3	0	1	4	8
110	2	4	1	3	10
111	2	3	0	2	7
112	0	0	2	2	4
113	2	1	1	1	5
114	3	2	1	1	7
115	2	4	3	4	13
116	1	0	0	0	1
117	2	2	3	2	9
118	0	0	1	1	2
119	0	1	1	0	2
120	3	0	1	1	5
121	1	1	3	0	5
122	0	0	1	0	1
123	3	1	1	0	5
124	1	0	0	1	2
125	0	0	2	0	2
126	2	0	1	0	3
127	0	0	0	0	0
128	0	0	0	0	0
over 128	7	7	2	2	18

The distributions were compared pairwise using the Kolmogorov-Smirnov Two Sample Test. The null hypothesis that there were no differences could not be rejected at a 10% level of significance.

The results from Table 2 show that almost all trials are within a one percent range from the lower bound (row 1) for $n \geq 10$. Note also that the entries in the first rows of tables 2 and 3 happen to increase as n increases.

REFERENCES

[1] Giglio, R. J. and H. M. Wagner, "Approximate Solutions to the Three Machine Scheduling Problem," Operations Research, *12*, 305–324 (1964).

[2] Johnson, S. M., "Optimal Two and Three-Stage Production Schedules with Setup Times Included," Naval Research Logistics Quarterly, *1*, 61–68 (1954).

[3] Szwarc, W., "Mathematical Aspects of the $3 \times n$ Job-Shop Sequencing Problem," Naval Research Logistics Quarterly, *21*, 145–153 (1974).

[4] Szwarc, W., "Special Cases of the Flow-Shop Problem," School of Business Administration, University of Wisconsin-Milwaukee, Dec. 1975; to appear, Naval Research Logistics Quarterly, *24*, No. 3, Sept. 1977.

A CONVEX PROPERTY OF AN ORDERED FLOW SHOP SEQUENCING PROBLEM*

S. S. Panwalkar

Department of Industrial Engineering
Texas Tech University
Lubbock, Texas

A. W. Khan

American Science and Engineering Company
Houston, Texas

ABSTRACT

A flow shop sequencing problem with ordered processing time matrices is considered. A convex property for the makespan sequences of such problems is discussed. On the basis of this property an efficient optimizing algorithm is presented. Although the proof of optimality has not been developed, several hundred problems were solved optimally with this procedure.

1. INTRODUCTION

Consider the classical n job m machine flow shop sequencing problem. Recently, Smith, *et al.*, [1] have introduced a subcategory of the flow shop problem with "ordered processing times." Two properties are stipulated for such an "ordered flow shop problem."

(i) If a particular job 'a' has a smaller value of processing time on any machine compared to another job 'b', then job 'a' will have smaller (or equal) processing time compared to job 'b' on all machines, and (ii) if any job has its j^{th} smallest processing time on any machine, then every job will have its j^{th} smallest processing time on the same machine.

Note that due to the characteristics described above, we can conveniently number jobs in the ascending order of processing times. In the following we will consider the ordered flow shop problem with minimum makespan criterion and permutation schedules only.

2. NOTATION

Let N denote the set of n jobs and let the jobs be numbered in ascending order of their processing times. Job n will then be referred to as the "largest job."

*Partial support for this research was provided by the National Science Foundation Grant GK–2869.

Let the set N be divided into two partitions, σ and σ' such that σ contains at least one job (σ' may be empty if $\sigma=N$). Assume for convenience that the largest job is always included in σ. Let r denote the number of jobs in $\sigma(1\leq r\leq n)$.

3. PREVIOUS RESULTS

Smith *et al.*, [1, 2] have presented some interesting results which can be represented by the following two theorems. The statements of the theorems are modified slightly to match the above notation and the proofs of these theorems may be found in [1, 2].

Theorem 1 [1]: In the ordered flow shop problem a minimum makespan sequence is given by arranging jobs in descending (ascending) order of processing times if the largest processing time for every job occurs on the first (last) machine.

THEOREM 2 [2]: In the ordered flow shop problem there exists a minimum makespan sequence of the form $\sigma\sigma'$, where jobs in σ are arranged in ascending order of processing times followed by jobs in σ' in descending order of processing times.

It may be noted that the second theorem is a more general one. Thus there are 2^{n-1} sequences (for all values of r) in which an optium sequence can be found. Each sequence satisfying Theorem 2 is said to have a "pyramid structure" and for a sequence with pyramid structure, the value of r indicates the position of the largest job. Smith *et al.*, [2] have presented an enumeration algorithm (S-P-D method) to evaluate 2^{n-1} sequences and recommend the use of a branch and bound procedure for further improvement.

4. CONVEX NATURE OF THE ORDERED PROBLEM

In order to explore some additional properties of the ordered problems (and possibly to develop a more efficient solution procedure), several problems were solved by the S-P-D method. For each problem, the makespan values of all 2^{n-1} sequences were analyzed.

For a given value of r defined above, there are ${}_{n-1}C_{r-1}$ possible sequences with a pyramid structure. Let S_r represent the set of these sequences. Let S^*_r represent the best sequence and T^*_r, the corresponding value of the makespan. If an optimal sequence for the complete problem (satisfying Theorem 2) occurs corresponding to $r=k$, it was observed that for all enumerated problems

$$T^*_1\geq T^*_2\geq \cdots \geq T^*_{k-1}\geq T^*_k\leq T^*_{k+1}\leq \cdots \leq T^*_n$$

This characteristic will be explained with an example. Consider a 6 job 6 machine problem in Table 1. For this problem the S-P-D algorithm will generate 32 sequences.

TABLE 1.—*Processing Time Matrix for an Ordered Flow Shop Problem*

Job	Machine					
	1	2	3	4	5	6
1	3	13	7	6	8	9
2	19	85	42	37	50	55
3	26	116	57	51	68	75
4	29	130	64	56	76	84
5	36	161	80	70	95	104
6	68	304	150	132	179	197

These sequences can be divided into the 6 sets as follows:

S_1=(654321)

S_2=(165432, 265431, 365421, 465321, 564321)

S_3=(126543, 136542, 146532, 156432, 236541, 246531, 256431, 346521, 356421, 456321)

S_4=(123654, 124653, 125643, 134652, 135642, 145632, 234651, 235641, 245631, 345621)

S_5=(123465, 123564, 124563, 134562, 234561)

S_6=(123456)

The sequence with minimum makespan in each of the above sets and the corresponding value of makespan are given below.

S^*_1=654321	T^*_1=1357
S^*_2=265431	T^*_2=1338
S^*_3=126543	T^*_3=1332
S^*_4=123654	T^*_4=1373
S^*_5=123465	T^*_5=1419
S^*_6=123456	T^*_6=1476

The makespan for the best sequence in each set is plotted in Figure 1 against the position of the largest job n in that sequence. With each change in position of the largest job starting with any end, the value of makespan decreases until the optimal value is reached and then it increases.

It may be noted that although the figure represents discrete quantities, the lines drawn through various points represent a "convex" shape. A mathematical proof for such a property has not been developed. All enumerated problems, however, exhibited this property.

A simple algorithm based on this property has been presented in the appendix. Two hundred problems ranging in sizes 6×15 to 14×15 were solved by the proposed method as well as a branch and bound method using Theorem 2 properties. All problems by the proposed method gave optimal results (as confirmed by the branch and bound solutions). The proposed method proved to be more efficient in all cases. Finally, we would like to encourage readers interested in the problem to develop mathematical proofs.

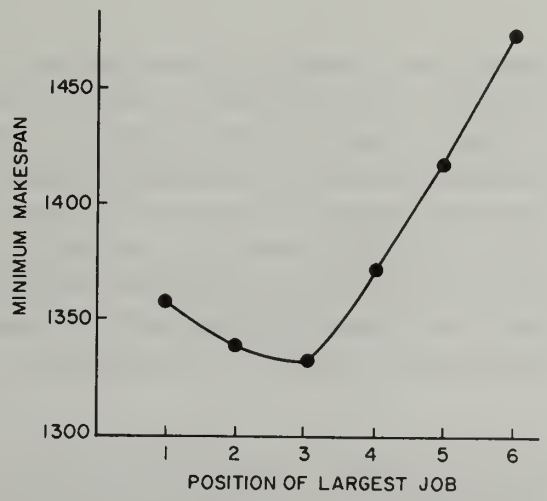


FIGURE 1.—Variation of minimum makespan with the position of largest job n.

APPENDIX

The step by step algorithmic procedure for the proposed algorithm is as follows:

STEP 1: Renumber the jobs according to the ascending order of magnitudes of the processing times.

STEP 2: Set $i=n+1$ and set $T^*_{n+1}=\infty$.

STEP 3: Set $i=i-1$. If $i=0$, go to Step 6.

STEP 4: With largest job in the i^{th} position evaluate all sequences with pyramid structure for makespan (i.e., all sequences with jobs in ascending and then descending order of processing times). When each new sequence is generated, compare with the previous best sequence and store only the best sequence. When all the pyramid shaped sequences with largest job in i^{th} position are evaluated, denote the best dequence by S^*_i.

STEP 5: If $T^*_i \leq T^*_{i+1}$ go to Step 3.

STEP 6: The sequence S^*_{i+1} is an optimal sequence.

REFERENCES

[1] Smith, M. L., S. S. Panwalkar, and R. A. Dudek, "Flow Shop Sequencing with ordered Processing Times Matrices," Management Science, *21*, 544–549 (1975).

[2] Smith, M. L., S. S. Panwalkar and R. A. Dudek, "Flowshop Sequencing Problem with Ordered Processing Time Matrices: A General Case," Naval Research Logistics Quarterly *23*, 481–486 (1976).

A MANPOWER PLANNING/CAPITAL BUDGETING MODEL (MAPCAB)

Rolf H. Clark

Office of Chief of Naval Operations
Washington, D.C.

Robert A. Comerford

University of Rhode Island
Kingston, Rhode Island

ABSTRACT

A deterministic resource allocation model is developed to optimize defense effectiveness subject to budget, manpower, and risk constraints. The model consists of two major submodels connected by a heuristic. The first is a mathematical program which optimizes the multiperiod weapon mix subject to the constraint set. The second is a manpower supply model based on a transition matrix in which individual transitions are functions of personnel related budgets and historical transition rates. The heuristic marries the submodels through an iterative process leading to improved solutions. An example is provided which demonstrates how systems are undercosted and overprocured if manpower supply is not properly reflected relative to manpower demand.

I. INTRODUCTION

Background

In economic decisions costs are normally assigned to an investment dependent on the inputs used to produce and operate it. Similarly, weapon systems are typically costed by the resources needed to develop, operate, and support them.* Because of the unique conditions of military manpower development, costing the manpower inputs as a function of manpower used is not only inaccurate, but actually erroneous. The error results because manpower levels required to man a set of weapons may require large pipeline inventories of skills not fully used, but needed to ensure adequate levels of more senior or more specialized skills which *are* used. If costs of a weapon system are determined by the manpower utilized directly, without accounting for the pipeline inventories, then the tendency will be to under-cost and over-invest in such systems. This will result in inability to man them adequately.

*As an example, the U.S. Navy's guidelines for economic analysis as stated in Secretary of the Navy Instruction 7000.14 of 14 March 1973 specify that personnel costs are charged according to ". . . the cost of military personnel services involved directly in the work performed."

Unaccounted for "pipeline" personnel costs apply not only to manpower waiting to be developed into more useful skills, but also to manpower which has transitted from a junior but valuable skill to a more senior but overmanned skill. Retired personnel are an obvious example, but other senior active skills may be less real worth (in terms of opportunity costs) than a more junior grade. The problem of improper weapon accumulation is compounded if improper discounting techniques make the present value of future pension costs less than they should be.* The result is to obtain systems more manpower intensive than is efficient.

An alternative approach avoiding the pipeline error is needed. Manpower supply must be properly accounted for in allocating Defense resources. In the proposed model, decision making in peacetime is the assumed scenario, though extensions to wartime are straightforward. The basic question will be what mix of weapon systems to accumulate over a long term planning horizon. The present value of the selected weapon configuration serves as the objective. The model accounts for all personnel types, including those such as recruits, trainees, and retirees, who are not specifically productive. It also can account for all budget allocations within the annual budget. These allocations are called subbudgets, and parallel actual Defense expenditures.

The model reflects the following assumptions:

1) Defense budgets are obligated annually. They are based on national macro-economic policy (which, in turn, reflects national security policy), and are functions principally of GNP and the unemployment rate. Since they are determined by internal factors, at least in peacetime, defense budgets are reasonably predictable.

2) Defense budgets are fixed for the year. Defense cannot borrow or lend funds to increase or diminish the annual budget.

3) While the social cost of capital is basic to the decision society makes when it decides how much to allocate to Defense, it is not equivalent to the cost of capital which Defense should base its decisions on once it has received the funds. That defense investment and the social cost of capital are not closely related has been upheld empirically. During periods of high world tension, defense expenditures have reflected imputed discount rates in excess of 20% while the social cost of capital (as measured by the treasury bond rate) remained at 3% during the same period.†

4) The military labor market is very imperfect. Capital is not a surrogate for manpower in the short run. Dollars cannot be converted into manpower for two reasons. First, military personnel are hired only as recruits, with needed skills then developed. Second, all equal grades are paid essentially the same wage regardless of relative worth. (The existence of specialty pay counters but does not neutralize this point.)

5) Due to the unforseeable tumultuous effects of sudden changes in Defense posture, changes in Defense policy must occur gradually to allow the economy to adjust. Also, risk considerations, such as maintaining some minimum defense posture, must be accommodated—and initial conditions of the analysis must match current defense configurations.

---

*Department of Defense Instruction 7043.3 of 18 October 1972, "Economic Analysis and Program Evaluation for Resource Management" specifies a 10% discount rate for all defense investment decisions. The proper rate for discounting social investments is discussed from a conceptual standpoint in Arrow and Kurz [1]. But as will be seen from the model below, an even more serious error than an improper rate is to disregard future pension costs altogether. This practice results from using limited planning horizons, and since pension costs are among the last to occur, causes manpower costs to be biased downward.

†For a discussion of discount rates in excess of 20% see Hitch and McKean [7, p. 211].

Rationale for the Manpower Planning/Capital Budgeting (MAPCAB) Model

Model development is based on the above assumptions. Collectively these assumptions make the Defense situation unique. However, some capital budgeting techniques borrowed from business decision theory, combined with familiar optimization methodology, can clarify the Defense resource allocation problem. MAPCAB combines some standard manpower planning techniques with capital budgeting criteria in a mathematical program to yield solutions. An iterative methodology used in the model makes the underlying interrelations between manpower and capital allocation visible to the decision maker. This facilitates coordination between the policy maker and the analyst when the model is parametrically exercised.

The model has the following features:

1) Superficially, the proposed technique for discounting is radically different from the current cost of capital orientation. The justification for the discount rate concept is, as usual, the relative advantage of current consumption over future consumption. But two questions arise. First, if Defense is to discount, what is the equivalent of "consumable" for defense? Second, what is the value of current consumables relative to future ones? Defense exists to deter, defend, or attack enemies. It can do that only with employable weapon units like naval task groups, ballistic missile submarines, bomber wings, missile defense units, etc. Resources can either be expended to have such weapon units now, or they can be directed toward building shipyards, conducting research and development, and building training facilities so that more weapon units will be available later. Viewed in this light, the consumables are employable weapon units. The worth of current relative to future weapons is obviously dependent on the tension level now versus the future. Consequently the objective function consists of the cumulative present value of deployable weapon systems, discounted as a function of the tension level. The tension level is measured as the probability of engagement in hostile activity during each year of the planning horizon. (The probabilities could be obtained from a panel of intelligence experts through Delphi techniques.*) The objective function variables can be adjusted for tactical and/or physical depreciation, reliability, and relative or absolute payoff.

2) The objective is maximized subject to budget and manpower constraints. Budgets and manpower limitations are estimated for a multiperiod planning horizon. Since future manpower supply is partially controlled by certain budget allocations (e.g. recruiting and training budgets), alterations to such budgets are made and new manpower estimates obtained. Changing manpower levels in turn affects budget allocations. In essence an entire new set of constraints is obtained. The new optimum objective is compared with the original solution. Through a process of such iterations, manpower supply affects weapon selection and weapon selection affects manpower supply, until a solution reasonably near an optimum obtains.

3) Large shifts in defense posture could cause economic tidal waves in the private sector. Such effects are prevented in the model by constraints which restrict year-to-year changes in weapon levels and/or budget allocations to a predetermined percentage of the earlier year's level. Additionally, risk is included as a constraint which places lower bounds on selected variables.

4) Finally, the model assumes divisibility in the variables. While one could argue for integral weapon units, most weapons are either sufficiently numerous to negate the value of an integer solution (e.g. FBM Submarines), or, in fact, are divisible (task groups).

---

*For a description of Delphi see Quade, "When Quantitative Models are Inadequate" [9, pp 333–343].

The MAPCAB model as presented in the next section incorporates the above factors in a linear program. It uses opportunity costs derived from the optimization process to make the budget allocation decisions which alter future manpower supply. The procedure compares normalized opportunity costs of the various scarce factors, and then heuristically alters manpower related budgets. Perceived shortages in manpower relative to capital are reduced by deducting funds from non-personnel related budgets and applying them to recruiting more people, or raising wages to retain people. Perceived shortages in skilled relative to unskilled manpower are reduced by shifting recruiting funds to training budgets. Budget reallocation continues so long as improvements in the objective function result. Eventually the opportunity costs will yield no further improvement within the budget flexibility allowed.

The model is intended to assist in long term policy decisions. Directions rather than exact magnitudes of movements are perhaps the most relevant output of a policy model. The type of policy questions which the model can answer include the following: Should manpower levels be increased or decreased? Should a particular system be expanded or phased out? What skills will be in the shortest supply if a certain mix is accumulated?

This initial model is entirely deterministic. Stochastic extensions accounting for distributions of annual budgets and manpower transitions are proper for follow on studies.

II. MODEL DETAILS

A. Notation

$\underline{A}$	Vector A.
$\underline{\underline{A}}$	Matrix A.
a_{kjt}	Amount of budget k needed per unit weapon j year t.
$\overline{B}$	Average value of budget dual variables for planning horizon T. (See equation (D–3).)
B_{kt}	Budget type k available in year t.
B_{rt}	Recruiting budget for year t.
B_{st}	Training budget for year t.
B^*_{tt}	Budget B_{tt} adjusted to reflect relative worth of money to manpower duals. (See equations D–4, D–6 and D–7).)
B^{**}_{tt}	Budget B_{tt} adjusted to reflect relative worths of money to manpower duals and relative worths of specialized to unspecialized skills. (See equations (D–12 and D–13).)
B_t	Total budget allocated to Defense in year t.
$c_{\underline{rgc}t}$	Accounting cost of skill code $\underline{rgc}$ in year t. Includes all personnel related costs such as wages, incentives, housing, medical, and administrative costs.
$C_{\underline{rgu},\underline{rgs}}$	The accounting cost of converting a person of skill code $\underline{rgu}$ into skill code $\underline{rgs}$.
D_t	The discount rate used in discounting the objective for year i. This rate is based on the world tension level for year i rather than on the cost of capital.
$d_{\underline{rgc}t}$	Dual variable associated with the constraint associated with the index $\underline{rgct}$.
$d'_{\underline{rgc}t}$	The duel variable d adjusted to dollar units. (See equation (D–14).)
$d''_{\underline{rgc}t}$	The adjusted dual d' adjusted a second time for the training cost of developing skil code $\underline{rgc}$. (See equation (D–15).)

$\overline{M}$ The average value of all duals associated with manpower constraints averaged over the entire planning horizon T. (Equation (D–2).)

m_{jt} The number of men of skill code $\underline{rgu_j}$ which can be converted to code $\underline{rgs_t}$ if the recommended budget amount is added to the current budget for such conversions. (See equation (D–20).)

M_{00t} The number of recruits obtained in year t.

M_{p0t} Number of retired persons living in year t.

M_{q0t} Number of persons deceasing in year t.

$\underline{M}_{rgc0}$ The vector of manpower levels existing at time 0.

m_{rgcjt} The number of men of skill code $\underline{rgc}$ required to man a unit of weapon type j in year t.

M_{rgct} The supply of men of type $\underline{rgc}$ available in year t.

$\underline{M}_{rgct}$ The vector of all manpower skill codes available in year t.

N_t Total number of dual variables associated with constraints indexed by manpower type $\underline{rgi}$ for the entire planning horizon T. If there were n different $\underline{rgi}$ type skills in each year t, then $N_t=nT$.

P_{jt} The relative payoff or utility of weapon type j in year t. P reflects reliability, depreciation, obsolescence, and mission weighting, and is discussed in formulas (B–1) and (B–4).

$\underline{rgc}$ The index indicating manpower type rate/grade/code rgc. The underlining of an index does not imply vector notation; rather it indicates a single index. The triple designation is used to facilitate implementation.

R_{kt} A multiplier obtained through the heuristic defined in formula (D–5), which denotes that portion of the subbudget B^*_{rt} to deduct from operating budget B_{kt} when adjusting budgets to provide more manpower. Equations (D–4) and (D–6) pertain.

R'_{kt} Analogous to R_{kt}, but designating the amount of increase required in budget B_{kt} to reduce manpower supply. Equations (D–7) and (D–8) pertain.

$\overline{S}$ The average value of all duals associated with specialized manpower constraints averaged over the entire planning horizon T. (See equation (D–9).)

$\overline{S}_t$ The average of all duals associated with specialized manpower constraints averaged over year t. Equation (D–11) pertains.

S_t The fraction of the reallocated recruiting budget which should be applied to developing specialty $\underline{rgs_t}$ in year t. (See formulas/(D–16), (D–17), and (D–20).)

T The planning horizon, ie. $t=0, 1, 2, \ldots, T$.

$\underline{\underline{T}}_t$ The transition matrix for year t characterized in C–1.

T^*_{ij} The readjusted transition rate reflecting the change in T_{ij} necessitated by the budget reallocations which effect manpower supply. (See equations (D–21) and D–22).)

U_j The heuristic denoting the fraction of the budget reallocation to develop new specialized skills $\underline{rgs_t}$ which should be used to retrain specifically skill $\underline{rgu_j}$. (See equations (D–19) and (D–20).)

$\overline{U}$ The average value of all duals associated with unspecialized manpower constraints averaged over the entire planning horizon T. (See equation (D–10).)

$\overline{U}_t$ The average of all duals associated with unspecialized manpower constraints averaged over year t. Equation (D–11) pertains.

x_{jt} The amount of type j weapon units to be deployable in year t.

Y_{rt} The amount by which the recruiting budget B^*_{rt} should be shifted to or from training budget B_{st} in year t. (See equations (D–11), (D–12) and (D–13).)

B. Optimization Phase

Objective Function

The objective function is assumed linear. Extensions to include diminishing marginal returns are appropriate but not considered here. Let

$$(x_{jt}: j=1, 2, \ldots, J; t=1, 2, \ldots, T)$$

represent the set of deployable weapon systems selected for the planning horizon T, where x_{jt} represents the number of units of weapon type j available in year t. Assume there are known factors r_{jt}, q_{jt}, and s_{jt} representing reliability, physical depreciation, and tactical depreciation, respectively. These three factors are included in this discussion because reliability, physical depreciation, and tactical depreciation are all parameters characterizing a weapon unit. A brief digression to discuss how values for each can be obtained is therefore appropriate. Reliability r_{jt} reflects, essentially, the "up time" of a weapon unit. If (at time t) it requires four aircraft in the inventory to keep one operational, then r_{jt} would be .25. Such data is available for most systems.

Physical depreciation q_{jt} represents system decay in the form of parts support. As systems age, repair and support costs tend to increase; thus q_{jt} would normally be decreasing in time. Again, data for such depreciation should be available from logistic records. A depreciation factor similar to accounting depreciation could be used to obtain values of q_{jt} for some systems if data is lacking.

Tactical depreciation is a function of the potential adversary's systems. Values for s_{jt} would have to be based on historical estimates of system tactical life. An estimate of tactical decay for model purposes might, if data is lacking, arbitrarily be based on exponential decay, with the decay rate being a function of the relative lifetime of different systems—e.g., aircraft being useful for 10 years, ships for 30, radars for 6, etc.

Assume now that there exist known measures of relative utility (in the foreseen scenarios) for the different weapon types. Denote these p_{jt}. These are of more theoretical impact to the model than the above three factors. The p_{jt} are the required output measures of the different weapon units during the different time frames—$p_{jt}x_{jt}$ would represent the raw output of the jth system in period t. How does one obtain the p_{jt}? Ideally, through a Delphi technique or other opinion sampling process, one attains them from military experts. As an initial estimate, however, one could let p_{jt} equal the unit cost of a system divided by the system's expected lifetime. This method implies obvious assumptions, one being that the cost of a system is a measure of its relative worth. This may not always be true, although it should be valid when a system is first produced. Nonetheless, this is an adequate starting point, one that lends itself to discussion by military planners, and therefore to subsequent adjustments to more accurate values of p_{jt}.

Of course, the aim is to obtain objective function coefficients reflecting the relative utilities of different systems at different times, adjusted for reliability, and for physical and tactical depreciation. Let these coefficients be denoted P_{jt}. Then

$$P_{jt}=f(p_{jt}, r_{,jt} q_{jt}, s_{jt}).$$

We will assume f takes the following specific form:

(B–1) $$P_{jt}=p_{jt}r_{jt}q_{jt}s_{jt}$$

The *unadjusted* effectiveness of all the weapon units available in year t, that is total defense system effectiveness in year t, then becomes

$$\sum_j P_{jt}x_{jt}$$

We would like to convert each year's defense capability to a "present value." If there exists some time preference for defense capabilities, then this can be done by an adjustment similar to time discounting. In fact, if D_t represents the discount rate in year t, then the present value of year t's output becomes

(B–2) $$\frac{1}{\prod_{i=1}^{t}(1+D_i)}\sum_j P_{jt}x_{jt}$$

The D_t should reflect the relative needs for defense systems at different times. This translates into having high values for D_t during high world tension levels, and low values during low tension levels. State Department and Defense Department planners might be able to predict relative tension levels—certainly in the past the buildups to wars have been quite obvious, and estimates on the return to peacetime have been made. The D_t's ideally would reflect these estimates. If agreement is not possible, then a constant rate for D_t, such as 5%, could be used. This is similar to procedures in *financial* discounting, where some constant rate (e.g., 10%) is used when agreement on the time structure of interest rates is not possible.

The overall present value of the objective function then is the present value of the set of selected systems x_{jt}. This is written:

(B–3) $$\text{Present Value}=\sum_t\left\{\frac{1}{\prod_{i=1}^{t}(1+D_i)}\sum_j P_{jt}x_{jt}\right\}$$

If mission versatility of the various systems is a factor, let s_{mj} be a measure of the effectiveness of weapon type j in mission m and let w_m be the weight specifying the relative importance of mission m in the overall defense scenario. The objective function than becomes

(B–4) $$\sum_t\left\{\frac{1}{\prod_{i=1}^{t}(1+D_i)}\sum_m w_m\sum_j P_{jt}s_{mj}x_{jt}\right\}$$

Hereafter expression (B–3) will be used to refer to the objective; maximization is understood.

Constraints

1) Budget constraints: Let a_{kjt} be the amount of budget type k needed to support one unit of weapon type j in year t. And let B_{kt} be the allocation of funds to subbudget k in year t. The B_{kt}; $k=1, 2, \ldots, T$, include only those subbudgets not directly dependent on manpower levels—(examples of included budgets are weapons procurement, support equipment, non-housing construction, and central supply. Excluded budgets are wages and incentives, personnel administrative support, medical, and housing.) The budget constraints are the following for the non-personnel related budgets:

Budget Constraints for Year t

$$\begin{aligned}
&\text{Budget 1: } a_{11t}x_{1t}+\ldots+a_{1jt}x_{jt}+\ldots+a_{1Jt}x_{Jt}\leq B_{1t}\\
&\qquad\vdots\qquad\qquad\qquad\vdots\\
&\text{Budget k: } a_{k1t}x_{1t}+\ldots+a_{kjt}x_{jt}+\ldots+a_{kJt}x_{Jt}\leq B_{kt}\\
&\qquad\vdots\qquad\qquad\qquad\vdots\\
&\text{Budget K: } a_{K1t}x_{1t}+\ldots+a_{Kjt}x_{jt}+\ldots+a_{KJt}x_{Jt}\leq B_{Kt}
\end{aligned}\tag{B–5}$$

There will be one such set of constraints for each year t. The personnel related budgets excluded from (B–5) enter the analysis through considerations of manpower supply below. Note the sum of the B_{kt} for each year t must be less than or equal to the annual budget B_t less the sum of all personnel budgets.

2) Manpower constraints: If $m_{\underline{rgc}jt}$ is the amount of men of skill code $\underline{rgc}$ (representing a rate/grade/specialty-code combination) required to man and support one unit of weapon type j in year t, and $M_{\underline{rgc}t}$ is the estimated manpower of type $\underline{rgc}$ supplied in year t, then for year t the manpower constraints will be the following:

Manpower Constraints for Year t

$$\begin{aligned}
&m_{\underline{1111}t}x_{1t}+\ldots+m_{\underline{111}jt}x_{jt}+\ldots+m_{\underline{111}Jt}x_{Jt}\leq M_{\underline{111}t}\\
&\qquad\vdots\qquad\qquad\qquad\vdots\\
&m_{\underline{rgc1}t}x_{1t}+\ldots+m_{\underline{rgc}jt}x_{jt}+\ldots+m_{\underline{rgc}Jt}x_{Jt}\leq M_{\underline{rgc}t}\\
&\qquad\vdots\qquad\qquad\qquad\vdots\\
&m_{\underline{RGC1}t}x_{1t}+\ldots+m_{\underline{RGC}jt}x_{jt}+\ldots+m_{\underline{RGC}Jt}x_{Jt}\leq M_{\underline{RGC}t}
\end{aligned}\tag{B–6}$$

Again, one such set of constraints applies for each year t.

3) Risk Constraints: For reasons of intuition, politics, or stability, constraints placing either lower limits on the weapon units for any year or upper and lower bounds on certain ratios of the variables may be included. Several types of such constraints are:

a) Lower Bound Constraints. These reflect the concern of military planners to allow dips in effectiveness to occur below some minimum readiness level. The requirement that total effectiveness in year $t+1$ be at least R times that in year t, where R is a constant near 1.0, would be met by the following constraint:

$$\sum_j P_{j,t+1}x_{j,t+1}\geq R\cdot\sum_j P_{jt}x_{jt}\tag{B–7}$$

If R' is a positive constant less than 1.0, the following constraint ensures that system j not be phased out too suddenly:

(B-8) $$P_{j,t+1}x_{j,t+1} \geq R'P_{jt}x_{jt}.$$

b) Upper and Lower Bounded Ratios. Military planners may feel that the relative ratios of different systems be kept within certain limits. This form of risk constraint can be met by the following condition:

(B-9) $$L \leq \frac{x_{jt}}{x_{it}} \leq U;\ L, U \text{ constants.}$$

4) Technological Constraints: It may be that factors besides budget limitations prevent the build up of certain systems. The following constraints are alternative ways of restricting the accumulation rate:

(B-10) $$\frac{x_{j,t+1}}{x_{jt}} \leq R;\ R > 1,$$

(B-11) $$x_{j,t+1} \leq x_{jt} + j^*;\ j^* \geq 0.$$

5) Initial Conditions: The starting point of the analysis requires that planners begin from current weapon configuration $x_{10}, \ldots, x_{j0}, \ldots, x_{J0}$, the weapon mix effective at $t=0$. This can be accomplished either by specifying the first period variables or allowing some flexibility in their values by constraints of the type

(B-12) $$ax_{j0} \leq x_{j1} \leq bx_{j0};\ 0 \leq a \leq 1 \leq b$$

Such requirements have certain real-world value. First, they are realistic, as resources are seldom flexible enough to achieve large short term changes. Second, the starting point of the analysis coincides with the situation faced by decision makers, who must decide on future courses of action given the current status. Analytic recommendations and decision alternatives are more likely to be compatible, and the gap between decision maker and analyst is narrowed. Finally, by starting at the current mix and restricting movements to relatively small but realistic magnitudes, the assumption of linearity in the model is less restrictive.

C. Manpower Supply Phase

Manpower Transition Matrix

1) Let $\underline{M}_{rgct}$ be the vector of all manpower skill codes $\underline{rgc}$ for all years t. Then $\underline{M}_{rgct}$ represents the right hand side of the manpower constraints (B-6) above, augmented with those manpower codes which are not used in any of the weapon systems in the weapon mix selected—such as recruits, trainees in training, and retired personnel. The manpower supply phase of MAPCAB shows how $\underline{M}_{rgct}$ is obtained and how it is modified through changes in recruiting budgets, training budgets, etc., to provide manpower supply most advantageous to planners within the flexibility allowed. The normal approach to estimating $\underline{M}$ is through a process involving a Markov transition matrix.*

*Such an approach is described by Forbes [5, pp. 93–113].

Since in MAPCAB the transition rates are altered through budget manipulation and since the initial "state vector" will be non-stochastic, the matrix will be referred to simply as the (deterministic) transition rate matrix $\underline{\underline{T}}_t$. To isolate the skill codes $\underline{rgc}$ which are actually used to man systems from those manpower codes which are necessary but not used, the following additional notation is introduced. Let

$M_{\underline{00}t}$=Number of recruits obtained in year t

$M_{\underline{p0}t}$=Number of retired people in system in year t

$M_{\underline{q0}t}$=Number leaving (quitting or deceasing) year t

$T_{\underline{rgc},\,\underline{rgc'}\,t}$=transition rate from $\underline{rgc}$ to $\underline{rgc'}$ in t.

The transition matrix $\underline{\underline{T}}_t$ can then be represented:

(C-1)

	$M_{\underline{rgc'}\,t}$	$M_{\underline{p0}t}$	$M_{\underline{q0}t}$
$M_{\underline{00}t-1}$	$\cdots\ T_{\underline{00},\,\underline{rgc}}\ \cdots$	0	0
$M_{\underline{rgc}t-1}$	$\cdots\ T_{\underline{rgc},\,\underline{rgc'}}\ \cdots$	$T_{\underline{rgc},\,\underline{p0}}$	$T_{\underline{rgc},\,\underline{q0}}$
$M_{\underline{p0}t-1}$	$\cdots\ 0\ \cdots$	$T_{\underline{p0},\,\underline{p0}}$	$T_{\underline{p0},\,\underline{q0}}$

There will be one such matrix for each year t, with different values for the transitions likely.

Given the matrices $\underline{\underline{T}}_t$, the initial manpower supply vector $\underline{M_{rgc0}}$, and the scalar recruiting quantities $M_{\underline{00}t}$ for $t=0, 1, \ldots, T-1$, the manpower supply vector $\underline{M_{rgct}}$ for any year t can be determined using the following formula, which is derived in the appendix:

$$\underline{M_{rgct}}=\prod_{i=1}^{t}\underline{\underline{T}}_t\underline{M_{rgc0}}+\sum_{j=1}^{t}\prod_{i=j}^{t}\underline{\underline{T}}_{t+1}\underline{M_{00j}} \tag{C-2}$$

Subsequently the transitions in $\underline{\underline{T}}_t$ will be discussed in more detail to show how each transition is obtained. It will be found that some transitions are the result of personal preferences on the part of the personnel, while others are the direct result of funds made available for specific training. For the former, planners must rely on historical data and estimated relationships between civilian/military advantages, while the latter are more directly and accurately obtained.

It should be pointed out that the $\underline{\underline{T}}_t$, $i>1$, are specified once given the initial manpower vector $\underline{M_{rgc0}}$, the vector of recruiting budgets over time, the vector of training budgets over time, and the initial transition matrix $\underline{\underline{T}}_0$. For, given the preferences of the individuals concerned and the known quantities of recruits, trainees, etc., each year, the fact that the matrix rows sum to one determine, under certain realistic assumptions, the remaining $T_{\underline{rgc},\,\underline{rgc}}$, (i.e., those besides the transitions to skills, assuming fixed proportions of transitions to voluntary positions).

D. The Optimization/Manpower Supply Relationship

The optimization process described above yields primal and dual solutions under proper conditions. The proposed improvement to the allocation process involves a heuristic based on the dual variables (opportunity costs) associated with the primal budget and manpower constraints. To describe the heuristic, further definitions and notation are needed.

Let the subscripts u and s denote unspecialized and specialized skills. Such specialized training is acquired through formal training programs, such as schooling, at significant cost. A person of rate/grade rg who has received formal training of type s would then be classified by skill code $\underline{rgs}$, while the same rate grade without specialty training would be designated skill code $\underline{rgu}$. This dichotimization of the previous subscript $\underline{rgc}$ is necessary to allow taking the training budgets properly into account. The transitions to unspecialized skills are caused by such relatively uncontrollable factors as time in rate, personal desires of individuals, and individual motivation. The transitions to specialty skills, however, are mainly a function of training funds available, and it is here that the decision makers have some real flexibility in manpower planning, for budget allocations can be altered to provide increased or decreased transition rates to the various specialty skills.

The general approach of the heuristic in taking this flexibility into account is the following. First, an overall measure of manpower value relative to capital value is obtained by an aggregate comparison of the opportunity costs associate with manpower versus budget constraints. A determination is then made to either increase funds allocated to recruiting or to decrease them—the former if manpower duals exceed budget duals indicating a greater need for men than money. Then an aggregate comparison is made between the specialized and unspecialized skill codes, again using the opportunity costs as the comparison factor, and funds are then shifted from recruiting to training, or vice versa, depending on the relative values of the skills. Finally a determination is made as to which specialty skills should be increased or decreased, again basing the decision on the relative opportunity costs. These steps are now described in more detail.

Details of the Heuristic

1) Redistributing operating and recruiting budgets: All budgets included in the optimization phase will be referred to as operating budgets for convenience. Let the accounting cost of skill $\underline{rgu}$ be $c_{\underline{rgu}}$ and for $\underline{rgs}$ be $c_{\underline{rgs}}$. These annual costs include wages and incentive pay, plus prorated medical, housing, and administrative costs relatable to personnel. Let $d_{\underline{rgut}}$ and $d_{\underline{rgst}}$ be the dual variables, or opportunity costs, associated with skill codes $\underline{rgu}$ and $\underline{rgs}$ in period t. Then

(D-1)
$$d'_{\underline{rgit}} = \frac{d_{\underline{rgit}}}{c_{\underline{rgit}}};\ i = u, s$$

is the opportunity cost per man converted to opportunity cost per dollar, and is then in the same units, and therefore directly comparable with, the opportunity costs associated with budget constraints. Define the aggregate manpower related opportunity cost to be the average of all such converted manpower duals for all periods t, i.e.

(D-2)
$$\overline{M} = \frac{1}{N_u + N_s} \cdot \sum_t \left(\sum_{\underline{rgu}} d'_{\underline{rgut}} + \sum_{\underline{rgs}} d'_{\underline{rgst}} \right),$$

where the symbol

$$\sum_{\underline{rgi}}$$

means the summation is to be taken over all skill codes $\underline{rgi}$ and N_t is the number of $\underline{rgi}$ skill codes that exist during the time span T, i.e. the number of $\underline{rgi}$ type duals. Similarly, for all budgets B_{kt}, define

(D-3) $$\overline{B}=\frac{1}{KT}\sum_t\sum_k d_{kt}.$$

Intuitively, if $\overline{M}/\overline{B}$ is greater than 1, then it would seem efficient to shift funds from operating budgets into manpower. Manpower can be increased by increasing recruiting rates or decreasing resignation rates. The former may be accomplished by increasing recruiting efforts and/or making service life more attractive. The latter may result from pay raises or other personnel incentives, such as increased medical and retirement benefits. For expository reasons, only the effects of alternate recruiting budgets are detailed here. Let

$$\frac{\overline{M}}{\overline{B}}=1+x;\ x\geq -1.$$

The quantity x is a measure of the amount of redistribution proposed. The heuristic to be used is to increase recruiting budgets B_{rt} by an amount $(x/z)B_{rt}$, with z an appropriate number limiting the percentage change in the budget to some reasonable amount. In order to keep the annual budget B_t unchanged, some or all of the B_{kt} must be decreased.

Specifically, for all t, increase B_{rt} to $B_{rt}{}^*$, where

(D-4) $$B^*{}_{rt}=B_{rt}+(x/z)B_{rt}.$$

To maintain the annual budget at the original level, those budgets currently least binding should be decreased, "least binding" meaning having the lowest opportunity costs. To do this, define

(D-5) $$R_{kt}=\frac{1=\dfrac{d_{kt}}{\sum_k d_{kt}}}{\sum_k\left\{1-\dfrac{d_{kt}}{\sum_k d_{kt}}\right\}}.$$

Note that

$$\sum_k R_{kt}=1,$$

and that for small d_{kt}, R_{kt} is large. Therefore the desired redistribution conditions are met by the heuristic

(D-6) $$B^*{}_{kt}=B_{kt}-R_{kt}B^*{}_{rt}.$$

This last expression is exemplary of the heuristic used whenever a set of budgets is to be collectively decreased by some known amount. Were it the case that a set of budgets was to be *increased*, the expression would be

(D-7) $$B^*{}_{kt}=B_{kt}+R'{}_{kt}B^{*\prime}{}_{tt},$$

where

(D-8) $$R'_{kt}=\frac{d_{kt}}{\sum_k d_{kt}}$$

and $B^{*\prime}{}_{it}$ is the amount the budgets were to be increased, the subscript i vice r indicating that it need not be the recruiting budget which is used as the base.

2) Redistributing recruiting and training budgets:† In this subsection the details of subbudget allocation to balance specialized and unspecialized manpower opportunity costs are presented. The case where specialized manpower is relatively more valuable is treated here. The opposite case is straightforward. Let

(D-9) $$\overline{S}=\frac{1}{N_s}\sum_t\sum_{\underline{rgs}}\left\{\frac{d_{\underline{rgs}t}}{c_{\underline{rgs}t}}\right\}$$

and

(D-10) $$\overline{U}=\frac{1}{N_u}\sum_t\sum_{\underline{rgu}}\left\{\frac{d_{\underline{rgu}t}}{c_{\underline{rgu}t}}\right\}.$$

$\overline{S}$ and $\overline{U}$ are average converted (to dollar unit) opportunity costs for specialized and unspecialized skill codes averaged over the entire planning horizon T. Assume $\overline{S}/\overline{U}=1+y$; $y\geq -1$. If $y>0$, then funds should be shifted from developing unspecialized skills into specialized skills. This can be accomplished by shifting recruiting funds into training funds, especially in those years when specialized skills are most needed. To do so let $\overline{S}_t$ and $\overline{U}_t$ be the average of the adjusted duals for specialized and unspecialized skills for just the year t. For each t, decrease $B^*{}_{rt}$ by Y_{rt}, where

(D-11) $$Y_{rt}=\frac{\overline{S}_t}{\overline{S}_t+\overline{U}_t}\cdot(yB^*{}_{rt})$$

and designate the new annual recruiting subbudgets by

(D-12) $$B^{**}{}_{rt}=B^*{}_{rt}-Y_{rt}.$$

To absorb this decrease, increase each annual training budget, denoted B_{st}, by this same amount, that is, let

(D-13) $$B^{**}{}_{st}=B_{st}+Y_{rt}$$

where again B^{**} is the new budget.

The phasing consideration mentioned in the previous footnote should not be overlooked. If the delay time required to convert skill *rgu* into *rgs* is very short, the above is valid. But if it takes, say, a year to turn a recruit into a useful labor unit, and a year to develop an unspecialized skill into a specialized one, then the budget redistributions must be made one, two, or even more years

†The current model stresses selecting budgets to alter and how much to alter them. The question of *when* to alter them is best handled through a systems dynamics approach. The authors are currently testing a feedback model which investigates this phasing problem.

prior to the year used for computing $\overline{S}_t$ and $\overline{U}_t$. In the illustration which follows later, a one year delay is assumed for converting recruits to productive labor units, and zero delay is assumed for converting u skills to s skills.

3) Readjusting the manpower transition rates: This section describes how Y_{rt} is allocated toward developing the skills $\underline{rgs}$. Two factors apply—the relative values of the specialties and the training costs associated with developing each specialty. The relative values of the specialties are reflected in the opportunity costs $d_{\underline{rgst}}$. Let the training cost of converting skill $\underline{rgu}$ into specialized skill $\underline{rgs}$ be denoted $C_{\underline{rgu,\,rgs}}$. These costs are assumed constant over time, although the extension to allow changes over time are straightforward. They will also be assumed constant over u, since a skill $\underline{rgs}$ is developed through formal means such as schooling, and costs would be equal for all trainees.

Therefore recalling the definition

$$d'_{\underline{rgst}}=\frac{d_{\underline{rgst}}}{c_{\underline{rgs}}}, \tag{D-14}$$

now define

$$d''_{\underline{rgst}}=\frac{d'_{\underline{rgst}}}{C_{\underline{rgu,\,rgs}}}. \tag{D-15}$$

Here d'' is a weighted measure of the various opportunity costs, the weighting now being formed with respect to the relative accounting costs of the personnel and of the skill transiton costs. A three stage approach is followed. First it is determined how much of the Y_{rt} should go toward training each specialty $\underline{rgs}$ in year t. Then it is determined how that training should be distributed among the $\underline{rgu}$ skills, that is which $\underline{rgu}$ codes can best afford to be converted into $\underline{rgs}$ codes. Finally the transition rates $T_{\underline{rgu,\,rgs}t}$ are adjusted to maintain consistency in the transition matrix $\underline{\underline{T}}_t$. These three phases are now discussed separately.

Phase I—Distributing the Y_{rt} among specialties: Let $\underline{rgs}_i$ be a specific skill code from the set of codes $\underline{rgs}$. Then let

$$S_i=\frac{d''_{\underline{rgs},t}}{\sum_i d_{\underline{rgs},t}}. \tag{D-16}$$

Then S_i represents the fraction of Y_{rt} to allocate toward the development of specialty $\underline{rgs}_i$. (Note again S_i sums to unity and a large d'' implies a large allocation to the manpower type it is associated with). The amount allocated to training skill codes $\underline{rgs}_i$ in year t is

$$S_i Y_{rt}. \tag{D-17}$$

Phase II—Determining which unspecialized skill codes to convert into $\underline{rgs}_i$ codes: The funds represented by expression (D–17) are spent in converting some $\underline{rgu}$ personnel into $\underline{rgs}_i$ personnel. The specific codes to convert will be those in excess supply, or having the lowest opportunity cost. Let $\underline{rgu}_j$ be a specific skill code from the set $\underline{rgu}$, where $\underline{rgu}$ is now limited to those unspecialized skill codes which can be converted to $\underline{rgs}_i$. If, analogously to (D–14),

$$d'_{\underline{rgut}}=\frac{d_{\underline{rgut}}}{d_{\underline{rgu}}}, \tag{D-18}$$

then a small d' indicates a good prospect for conversion to $\underline{rgs}_i$. Therefore let

$$U_j = \frac{1 - \dfrac{d'_{\underline{rgu}_j}}{\sum_j d'_{\underline{rgu}_j}}}{\sum_j \left\{ 1 - \dfrac{d'_{\underline{rgu}_j}}{\sum_j d'_{\underline{rgu}_j}} \right\}}, \tag{D-19}$$

and allocate U_j times the total added expenditure on obtaining new $\underline{rgs}_i$ skills on converting $\underline{rgu}_j$ personnel into $\underline{rgs}_i$ personnel—that is allocate $U_j S_i Y_{ri}$ toward converting $\underline{rgu}_j$ into $\underline{rgs}_i$ (in addition to what was previously spent on the same conversion).

Phase III—Adjusting the transition rates: Each allocation $U_j S_i Y_{ri}$ can convert

$$\frac{U_j S_i Y_{ri}}{C_{\underline{rgu},\,\underline{rgs}_i}} = m_{ji} \tag{D-20}$$

men of skill code $\underline{rgu}_j$ into $\underline{rgs}_i$. Therefore the transition rate $T_{\underline{rgu}_j \underline{rgs}_i}$ for year t must be increased to reflect this additional conversion, and the remaining transitions from $\underline{rgu}_j$ into other skill codes (specialized or unspecialized) must be adjusted to ensure the sum of the rows of $\underline{\underline{T}}_t$ equal unity. Since it can be assumed that the $\underline{rgu}_j$ personnel being converted would otherwise have transitted to other unspecialized skills (in fact, most would have stayed $\underline{rgu}_j$ or advanced a grade to $\underline{r, g+1, u_j}$) the entire adjustment would be absorbed if the following facts are utilized. First the new transition rate $T_{\underline{rgu}_j \underline{rgs}_i}$ must be the ratio of the total number of $\underline{rgu}_j$ to $\underline{rgs}_i$ conversions performed in year t divided by the total manpower of type $\underline{rgu}_j$ available at the end of year $t-1$. Therefore the new transition rate is

$$T^*_{\underline{rgu}_j \underline{rgs}_i t} = T_{\underline{rgu}_j \underline{rgs}_i t} + \frac{m_{ji}}{M_{\underline{rgu}_j t-1}}. \tag{D-21}$$

The new transition rate from $\underline{rgu}_j$ to $\underline{rgu}_k$, where k indexes the unspecialized skills which can be converted into $\underline{rgs}_i$, is not usually controllable, (since persons choose to become specialized, rather than are forced to do so), and the most reasonable assumption is that the conversions come at the cost of $\underline{rgu}_k$ skills in the same ratios as the previous conversions to such skills. More specifically,

$$T^*_{\underline{rgu}_j \underline{rgu}_k t} = T_{\underline{rgu}_j \underline{rgu}_k t} - \left\{ \frac{T_{jk}}{\sum_k T_{jk}} \cdot \frac{m_{ji}}{M_{\underline{rgu}_j t-1}} \right\}, \tag{D-22}$$

where, for convenience,

$$T_{jk} = T_{\underline{rgu}_j \underline{rgu}_k t}.$$

E. The Iterative Process

The heuristics developed in the last section allow obtaining new transition matrices T^*_t which, using the time zero manpower supply vector $M_{rg:0}$ and the new estimated numbers of recruits for each year t, yield new right hand sides for the optimization phase.† These new right hand sides

†Equation (C-2) yields the new manpower levels.

should yield an improved solution without violating the annual budget limitations. If, on rerunning the optimization phase with these new constraints the solution becomes infeasible, the iteration should be repeated with each reallocation indicated by the heuristics decreased by a fraction, such as halved. If the solution results in a decrease in the objective, which may happen if the reallocations are so large that the optimum is passed, the reallocations should also be halved. If the steps are small enough in each iteration, the objective should show small improvements each iteration, until no further significant improvement is obtainable. This point will naturally be reached if the dual variables are approximately equal, or the limitations of the problem flexibility in terms of budget changes or manpower development is reached. Undesired occurrences, such as cycling or asymptotic behavior without significant improvement, should be easily analyzed due to the sequential nature of operations. This would not be the case if the problem was converted to a one phase optimization program.

The iterative scheme used here is a modification of the decomposition technique formulated by Kornai and Liptak in [8]. Computational forms of large scale applications of the above manpower planning model can benefit from the algorithm they originated.

F. An Illustration

Assume a hypothetical defense structure having three types of weapon units denoted A, B, and C. Also assume a four year planning horizon and a simple manpower structure consisting only of recruits, unspecialized skills u, and specialized skills s. The annual defense budget B_t is constant over time. The annual discount rates to be used are obtained from intelligence experts (if war is imminent, D_t is high; for distant years it could stabilize at, say, 5 percent). Also, the relative worth of the different weapon types is estimated by experienced military planners, as discussed.

Known Values:

- Annual budget=4,851 billion dollars.
- Relative net payoffs (P_{jt}): $A:B:C=100:10:25$
- Discount rate: $D_t=(0.10, 0.08, 0.05, 0.05)$. (Reflecting decreasing tension levels.)
- System requirements per unit:

	A	B	C
Operating budget (billions)	0.16	0.0183	0.0
Manpower type u (thousands)	4.0	0.3	0.655
Manpower type s (thousands)	1.9	0.16	0.3

Cost of training skill code s: \$3030 per man.

	recruit	u	s
Accounting cost per man:	\$5000	\$8000	\$10000

- Planned budget allocations:

	Year			
	1	2	3	4
recruiting (\$ billions)	0.045	0.045	0.045	0.045
training (\$ billions)	0.045	0.045	0.045	0.045

- Recruiting cost: $1000 per recruit.
- Trainees: From training budget and cost per trainee 14,850 s-type personnel are trained each year. Initially 4,500 of these come from recruit types and 10,350 come from u types.
- Initial transition matrix T_1:

	u code	s code	resigned
recruit	0.9	0.1	0
u code	0.35	0.15	0.5
s code	0	0.5	0.5

- Initial values:
  recruits=45,000
  u code =69,000
  s code =34,000
  weapon units:
    type A=10
    type B=50
    type C=20
- Restrictions: No weapon system can be decreased by more than 10% in any year.

Problems to be Considered:

With the above facts two problems will be solved. In the first, manpower constraints will be included in the operating budget rather than accounted for separately. This is the normal approach, with each system being charged for the manpower required to man and support it. The second problem, using the same data but explicitly accounting for manpower constraints, is solved, and a totally different solution, resulting in a completely different allocation of resources, results. The second problem is then revised through the heuristic rationale discussed to first yield an improved solution, and then to show what happens if the reallocation is carried too far. The data for these problems are not precise, but reasonably approximate figures of actual weapons used in the U.S. Defense arsenal to show the impact of using the wrong rationale.

Solutions:

Using formula (B–3) the objective function to be maximized is

$$\text{(G–1)}\quad 90.9x_{A1}+9.09x_{B1}+22.7x_{C1}+84x_{A2}+8.4x_{B2}+21.01x_{C2}$$
$$+80x_{A3}+8x_{B3}+20x_{C3}+76.3x_{A4}+7.63x_{B4}+19.08x_{C4}.$$

The manpower supply of codes u and s, using the initially planned constant recruiting and training budgets of 45 million dollars per year are (in thousands of men)

(G–2)

	Year			
	1	2	3	4
u code	64.65	63.13	61.96	60.95
s code	31.85	30.12	29.66	29.86

PROBLEM 1: Systems costed out with manpower implicit in operating costs.

The total annual budget of 4.851 billion dollars must be adjusted by deducting the annual training budget, the annual recruiting budget, and the annual wages paid to recruits who do not man or support any system. (The same would be true of retirees, who do not exist in this problem.) The available budget then is

$$4.851-0.090-(45{,}000\times\text{recruit wage})=4.536 \text{ billion dollars.}$$

Since manpower costs are to be included in the system operating costs, the budget factors for each weapon unit must be adjusted. For example, system A costs must be increased on a per unit basis by the cost of paying the skills used by the system, 4000 men times \$8000 or .032 billion for the u coded personnel used, and 1900 men times \$10,000 for the s coded personnel used. The per unit cost of system A is then the original .16 plus .032 plus .019 billion dollars. The new per unit costs are

$$\text{system A}=0.211$$

$$\text{system B}=0.0223$$

$$\text{system C}=0.0628$$

The appropriate budget constraints for years 1 through 4 thus are

(G–3) $$0.211x_{At}+0.0223x_{Bt}+0.0628x_{Ct}\leq 4.536;\ t=1,\ \ldots\ 4$$

The restriction that no system be depleted by more than 10% in any year is represented by

$$x_{jt}\geq 0.9x_{jt-1}$$

or

$$0.9x_{jt-1}-1x_{jt}\leq 0;\ j=A,\ B,\ C$$

$$t=2,\ 3,\ 4$$

and

$$x_{A1}\geq 8;\ x_{B1}\geq 45;\ x_{C1}\geq 18.$$

The optimum solution is presented in Table 3 for comparison with subsequent solutions. The matrix representation of the problem is presented in Table 1.

TABLE 1.—*Problem 1*

Variables:

x_{A1}	x_{B1}	x_{C1}	x_{A2}	x_{B2}	x_{C2}	x_{A3}	x_{B3}	x_{C3}	x_{A4}	x_{B4}	x_{C4}		
Objective: Maximize													
90. 9	9. 09	22. 7	84.	8. 4	21.	80.	8. 0	20.	76. 3	7. 63	19. 1		
Subject to:													
0. 21	0. 023	0. 068										$\leq$	4. 536
			0. 21	0. 023	0. 068							$\leq$	4. 536
						0. 21	0. 023	0. 068				$\leq$	4. 536
									0. 21	0. 023	0. 068	$\leq$	4. 536
. 9			−1.									$\leq$	0
	. 9			−1.								$\leq$	0
		. 9			−1.							$\leq$	0
			. 9			−1.						$\leq$	0
				. 9			−1.					$\leq$	0
					. 9			−1.				$\leq$	0
						. 9			−1.			$\leq$	0
							. 9			−1.		$\leq$	0
								. 9			−1.	$\leq$	0
1.												$\leq$	9. 0
	1.											$\leq$	45.
		1.										$\leq$	18.

PROBLEM 2: Manpower supply explicitly accounted for.*

In this case the annual budget constraints must be adjusted to reflect not only the recruiting, training, and recruit wage costs, but also for all other wages paid. Doing this yields budget constraints of 3.7, 3.73, 3.744, and 3.75 for years 1 to 4. The explicit budget and manpower constraints are shown in Table 2, and the solution is in Table 3. The important thing to notice is that the Problem 1 solution calls for increasing system A over time and decreasing system C, while Problem 2 calls for the opposite. This disparity results because in Solution 1, since manpower and money are treated interchangeably, the systems selected are those with a low per unit operating cost (per unit of payoff). But while the solution is feasible in terms of money, note it is infeasible in terms of manpower—for in year 1, as an example, the number of s personnel to man the selected systems are $11.38(1.9)+45(0.16)+18(0.3)=34.222$ thousand men. Yet the total available supply of such men is only 31,850. If the solution is accepted, and the systems bought, they will be undermanned, not an uncommon defense occurrence.

*The three phases of Problem 2 are represented by the constraints of Table 2.

TABLE 2.—*Problem 2*

Variables:

x_{A1}	x_{B1}	x_{C1}	x_{A2}	x_{B2}	x_{C2}	x_{A3}	x_{B3}	x_{C3}	x_{A4}	x_{B4}	x_{C4}

Objective: Maximize

90. 9	9. 09	22. 7	84.	8. 4	21.	80.	8. 0	20.	76. 3	7. 63	19. 1

Subject to

0. 16	. 018	. 055										$\leq$	B_1
4. 0	0. 30	0. 60										$\leq$	M_{u1}
1. 9	0. 16	0. 30										$\leq$	M_{s1}
			0. 16	0. 018	. 055							$\leq$	B_2
			4. 0	0. 30	0. 60							$\leq$	M_{u2}
			1. 9	0. 16	0. 30							$\leq$	M_{s2}
						0. 16	. 018	. 055				$\leq$	B_3
						4. 0	0. 30	0. 60				$\leq$	M_{u3}
						1. 9	0. 16	0. 30				$\leq$	M_{s3}
									0. 16	0. 018	0. 055	$\leq$	B_4
									4. 0	0. 30	0. 60	$\leq$	M_{u4}
									1. 9	0. 16	0. 30	$\leq$	M_{s4}
0. 9			−1.									$\leq$	0
	0. 9			−1.								$\leq$	0
		0. 9			−1.							$\leq$	0
			0. 9			−1.						$\leq$	0
				0. 9			−1.					$\leq$	0
					0. 9			−1.				$\leq$	0
						0. 9			−1.			$\leq$	0
							0. 9			−1.		$\leq$	0
								0. 9			−1.	$\leq$	0
1. 0												$\geq$	9. 0
	1. 0											$\geq$	45. 0
		1. 0										$\geq$	18. 0

Right Hand Sides:
(Constraints 1–12)

Designation	Prob 2	Iter 1	Iter 2
B_1	3. 7	3. 528	3. 474
M_{u1}	64. 65	68. 14	71. 6
M_{s1}	31. 85	38. 36	40. 4
B_2	3. 73	3. 51	3. 453
M_{u2}	63. 13	67. 87	69. 66
M_{s2}	30. 12	40. 38	41. 34
B_3	3. 744	3. 50	3. 411
M_{u3}	61. 96	67. 72	73. 9
M_{s3}	29. 66	4. 14	42. 1
B_4	3. 75	3. 496	3. 415
M_{u4}	60. 95	67. 68	73. 05
M_{s4}	29. 86	41. 89	42. 45

PROBLEM 2, ITERATION 1

If the dual variables of the solution to Problem 2 are investigated, it becomes apparent that manpower is more valuable than money. Furthermore, s skills are more valuable than u skills. If, therefore, the recruiting budget is first increased by 67% each year, and 67% of this increase is directed toward converting more recruits to s skills, the resulting manpower levels and cost changes result in the solution shown in Table 3. Notice that the objective value has increased, but no annual budget was changed. Further refinements are possible, an obvious one being to shift some training funds into recruiting since of the manpower related duals, only the u skills are now binding

TABLE 3.—*Solutions*

		Problem 1	*Problem 2*	*Iter 1*	*Iter 2*
Objective		6800	6430	6536	6450
Primal Variables	x_{A1}	11. 38	9	10. 71	10. 38
	x_{B1}	45	45	45	45
	x_{C1}	18	25. 17	18	18
	x_{A2}	12. 4	8. 1	11	11. 3
	x_{B2}	40. 5	40. 5	46. 5	40. 5
	x_{C2}	16. 2	27. 5	16. 2	16. 2
	x_{A3}	13. 3	7. 29	10. 7	12. 14
	x_{B3}	36. 4	36. 45	53. 8	36. 45
	x_{C3}	14. 6	33. 25	14. 6	14. 58
	x_{A4}	14. 1	6. 56	10. 4	13. 08
	x_{B4}	32. 8	44. 29	60. 7	32. 8
	x_{C4}	13. 1	34. 36	13. 12	13. 12
Dual Variables	Constraint				
	1		0	568	568
	2	Not	0	0	0
	3	Applicable	75. 7	0	0
	4		0	333	525
	5		0	7. 67	0
	6		70	0	0
	7		0	3. 7	500
	8		0	7. 3	0
	9		66. 7	0	0
	10		230. 7	302. 8	477
	11		0	6. 96	0
	12		21. 29	0	0

PROBLEM 2, ITERATION 2.

If the reallocations of budgets are carried too far, the increase in the objective may be less than for smaller readjustments. If the Problem 2 adjustments are 73% instead of the 67% made in Iteration 1, then the results are as shown in Table 3. The objective is less than for Iteration 1, and note that only the budget constraints are binding, as all other duals are zero. Manpower is now less valuable than money, and funds should be shifted into operating budgets and out of recruiting/training.

III. APPENDIX

Derivation of Equation (C–2):

Were it not for the scalar M_{00t} denoting the new recruits each year, the manpower vector required for the right hand side of the manpower constraints, $\underline{M}_{rgct}$, would be obtained as follows:

$$\underline{M}_{rgct} = \prod_{i=1}^{t} \underline{\underline{T}}_i \underline{M}_{rgc0}. \dagger \tag{III–1}$$

But this expression must be modified since the recruit-to-recruit transition rate is zero and new recruits arrive each year. That equation (C–2) holds will be shown in an inductive proof. First, for year 1, by definition

$$\underline{M}_{rgc1} = \underline{\underline{T}}_1 \underline{M}_{rgc0}.$$

Since $T_{\underline{00},\underline{00}t}$, the transition of recruits to recruits is zero for each year t, $\underline{M}_{rgc1}$ will have its first element equal to zero. To obtain a valid manpower supply vector in year 1 the recruits arriving in year 1 must be added to $\underline{M}_{rgc1}$. Define $\underline{M}_{00t}$, to be the column vector

$$(\underline{M}_{00t}, 0, 0, \ldots, 0),$$

and let

$$\underline{M'}_{rgct} = \underline{M}_{rgct} + \underline{M}_{00t}; \text{ for all } t.$$

For year 1,

$$\underline{M'}_{rgc1} = \underline{M}_{rgc1} + \underline{M}_{001},$$

And, therefore,

$$\underline{M'}_{rgc1} = \underline{\underline{T}}_1 \underline{M}_{rgc0} + \underline{M}_{001}.$$

---

†Since $\underline{M}_{rgc1} = T1\,\underline{M}_{rgc0}$ and $\underline{M}_{rgci} = \underline{\underline{T}}_i \underline{M}_{rgci-1}$, for all i from 2 to T, $(I-1)$ is true by induction. Also, T does not include the final column of (C−1), nor does $\underline{M}$ include the number of resignees since resignees are external to the system.

The next year's manpower vector can now be obtained:

$$\begin{aligned}\underline{M}_{rgc2} &= \underline{\underline{T}}_2 \underline{M}'_{rgc1} \\ &= \underline{\underline{T}}_2(\underline{\underline{T}}_1 \underline{M}_{rgc0} + \underline{M}_{001}) \\ &= \underline{\underline{T}}_2 \underline{\underline{T}}_1 \underline{M}_{rgc0} + \underline{\underline{T}}_2 \underline{M}_{001},\end{aligned}$$

and

$$\begin{aligned}\underline{M}'_{rgc2} &= \underline{M}_{rgc2} + \underline{M}_{002} \\ &= \underline{\underline{T}}_2 \underline{\underline{T}}_1 \underline{M}_{rgc0} + \underline{\underline{T}}_2 \underline{M}_{001} + \underline{M}_{002}.\end{aligned}$$

Proceeding in similar fashion,

$$\underline{M}_{rgc3} = \underline{\underline{T}}_3 \underline{\underline{T}}_2 \underline{\underline{T}}_1 \underline{M}_{rgc0} + \underline{\underline{T}}_3 \underline{\underline{T}}_2 \underline{M}_{001} + \underline{\underline{T}}_3 \underline{M}_{002},$$

and

$$\underline{M}'_{rgc3} = \underline{M}_{rgc3} + \underline{M}_{003}.$$

It is clear by continuing that the expression for $\underline{M}_{rgct}$ given in (C–2) is valid for the general case. The total manpower supply vector is then $\underline{M}_{rgct}$ augmented by the number of recruits entering in year t.

BIBLIOGRAPHY

[1] Arrow, Kenneth J., and Mordecai Kurz, *Public Investment, The Rate of Return, and Optimal Fiscal Policy.* (Baltimore: The Johns Hopkins Press, 1970).

[2] Carleton, Willard T., "Linear Programming and Capital Budgeting Models: A New Interpretation," The Journal of Finance, 825–833, (December, 1969).

[3] Charnes, A., and W. W. Cooper, *Management Models and Industrial Applications of Linear Programming* (New York: John Wiley & Sons, 1961).

[4] Charnes, A., W. W. Cooper and R. J. Niehaus, *Studies in Civilian Manpower Planning* (NAVSO P–3540) (Washington, D.C.: Navy Department, Office of Civilian Manpower Planning, July 1972).

[5] Forbes, A. F., "Markov Chain Models for Manpower Systems," In *Manpower and Management Science*, pp. 93–113. Edited by D. H. Bartholomew and A. R. Smith (Lexington, Mass.: D. C. Heath & Co., 1971).

[6] Hirshleifer, J., "On the Theory of Optimal Investment Decisions," Journal of Political Economy, 329–352 (August, 1958).

[7] Hitch, C. J., and R. N. McKean, *The Economics of Defense in the Nuclear Age* (Santa Monica, Calif.: The RAND Corporation, 1960).

[8] Kornai, J. and T. Liptak, "Two Level Planning," Econometricas, *33*, 141–169 (1965).

[9] Quade, E. S., and W. Boucher, eds. *Systems Analysis and Policy Planning—Applications in Defense* (New York: Elsevier, 1968).

[10] Weingartner, H. Martin, *Mathematical Programming and the Analysis of Capital Budgeting Problems* (Chicago: Markham Publishing Co., 1967).

AN F APPROXIMATION FOR TWO-PARAMETER WEIBULL AND LOGNORMAL TOLERANCE BOUNDS BASED ON POSSIBLY CENSORED DATA*

Nancy R. Mann

Rockwell International
Thousand Oaks, California

ABSTRACT

An approximation suggested in Mann, Schafer and Singpurwalla [18] for obtaining small-sample tolerance bounds based on possibly censored two-parameter Weibull and lognormal samples is investigated. The tolerance bounds obtained are those that effectively make most efficient use of sample data. Values based on the approximation are compared with some available exact values and shown to be in surprisingly good agreement, even in certain cases in which sample sizes are very small or censoring is extensive. Ranges over which error in the approximation is less than about 1 or 2 percent are determined. The investigation of the precision of the approximation extends results of Lawless [8], who considered large-sample maximum-likelihood estimates of parameters as the basis for approximate 95 percent Weibull tolerance bounds obtained by the general approach described in [18]. For Weibull (or extreme-value) data the approximation is particularly useful when sample sizes are moderately large (more than 25), but not large enough (well over 100 for severely censored data) for asymptotic normality of estimators to apply. For such cases simplified efficient linear estimates or maximum-likelihood estimates may be used to obtain the approximate tolerance bounds. For lognormal censored data, best linear unbiased estimates may be used, or any efficient unbiased estimators for which variances and covariances are known as functions of the square of the distribution variance.

1. INTRODUCTION

In the following we consider tolerance bounds for two-parameter Weibull and lognormal distributions, or equivalently, extreme-value (also known as Gumbel) and normal distributions, respectively. In the context of reliability, when the distributions are failure-time populations, then the tolerance bounds are confidence bounds on reliable life t_P, a population percentile corresponding to a specified survival probability $100R$ and $P=1-R$.

We assume that there may be type-II censoring of the data. In a life-testing context, this means that the life test applied to a sample of size n is terminated at the time of the rth failure, $r \leq n$. We order the observable variates from smallest to largest and call their logarithms $X_{1,n}, \ldots,$

*The research documented herein was supported by the Air Force Office of Scientific Research, AFSC, USAF under Contract No. F44620-71-C-0029.

$X_{r,n}$. Thus if the data are Weibull (lognormal), the sample of unordered X's is a sample of variates from the extreme-value (normal) distribution. In either case the distribution of the X's has a location parameter μ and a scale parameter σ. If the X's are extreme-value variates, their density function has the form

$$f_X(x)=\frac{1}{\sigma}\exp\left[-\exp\left(\frac{x-\mu}{\sigma}\right)\right]\exp\left(\frac{x-\mu}{\sigma}\right) \tag{1.1}$$

If the X's are normal, their density function has the familiar form

$$\phi_X(x)=\frac{1}{\sqrt{2\pi}\,\sigma}\exp\left[-\left(\frac{x-\mu}{\sqrt{2}\,\sigma}\right)^2\right] \tag{1.2}$$

For Weibull or extreme-value data, if censoring is not extensive, one can use Monte Carlo-generated tables of Thoman, Bain and Antle [22] and Billman, Antle and Bain [2] to obtain tolerance bounds. The tables are used in conjunction with iteratively obtained maximum-likelihood estimates of distribution parameters. If sample size n is 25 or less, one can determine tolerance bounds from tables of Mann and Fertig [14], Mann, Fertig and Scheuer [16], or Mann, Schafer and Singpurwalla [17] for $r=3(1)n$. These tables have also been generated by Monte Carlo procedures. They require the use of best linear invariant estimates

$$\sum_{i=1}^{r} a_{i,r,n}x_{i,n} \text{ and } \sum_{i=1}^{r} c_{i,r,n}x_{i,n}$$

of the parameters μ and σ, with values of $\{a_{i,r,n}\}$ and $\{c_{i,r,n}\}$ available in Mann [11, 12] and in [18]. Results of Engelhardt and Bain [4], Mann and Fertig [15], and Thoman, Bain and Antle [22] indicate that maximum-likelihood estimators and best linear invariant estimators give very nearly the same results for the extreme-value distribution, even for small sample sizes and rather extreme censoring. Moreover, for extreme-value data, the distributions of these two types of parameter estimators are very nearly the same.

For lognormal data that are not censored, tolerance bounds can be determined by simply disregarding the order of the observations, $x_{1,n}, \ldots, x_{n,n}$, and calculating $\bar{x}$ and s to be used with tables of the noncentral t-distribution.

The approximation to be described in Section 2 is suggested in [18] as a method for obtaining tolerance bounds under conditions of sample size or censoring for which the Monte Carlo generated tables or the noncentral t-tables described above are not applicable. The approximation is based on efficient linear estimators such as best linear invariant (B.L.I.) estimators or linear transformations of these estimators which are best linear unbiased (B.L.U.). One can also use maximum-likelihood (M.L.) estimators or efficient approximations to the optimum linear estimators, as long as information concerning variances and covariances of unbiased versions of the estimators is available. All of the B.L.I., B.L.U. and M.L. estimators contain essentially all the information in the sample that can be used in making inferences about the distribution parameters (see, for example, Lawless [7]). The approximate lower tolerance bounds described below are thus essentially the most accurate available (have, in effect, the minimum probability of falling below any percentile less than the true percentile of interest) among bounds independent of the parameters (see Lehman [14, p. 78]).

2. THE F APPROXIMATION

The result that provides the basis for the F approximation is that of Pyke [20] applying to any difference of adjacent ordered observable variates (called a "spacing" by Pyke) from a continuous distribution. Pyke showed that as sample size increases, the distribution of each such difference approaches that of a weighted exponential variate, or equivalently, that of a weighted chi-square with 2 degrees of freedom. Recently van Montfort [23] observed that any spacing

$$H_i = X_{i+1,m} - X_{i,m},\ i=1, \ldots, n, \tag{2.1}$$

from the extreme-value distribution, when divided by its expectation, has approximately an exponential distribution with mean 1 and variance very near to 1 [thus $2H_i/E(H_i)$ is approximately a chi-square with 2 degrees of freedom] *even for sample size* n *as small as 3*. Also, van Montfort observed that for a size-n sample of extreme-value variates with n as small as 3, the covariance between H_i and H_j, $i \neq j$, is approximately zero. Pyke [20] showed that in general H_i and H_j are *asymptotically independent* for $p_1 = i/n$ and $p_2 = j/n$ fixed with increasing n. Using tables, in Sarhan and Greenberg [21], of expectations, variances and covariances of reduced order statistics from the normal distribution, one can see that the properties observed by van Montfort for the extreme-value distribution hold for Gaussian spacings for sample sizes about 6 or larger.

Mann and Fertig [15] combined these results with that of Box [3], demonstrating that linear combinations of chi-squares are weighted chi-squares. They showed that, for sample size greater than about 3, any efficient unbiased linear estimator $\sigma^*_{r,n}$ of the extreme-value scale parameter with variance $C_{r,n}\sigma^2$ is such that $\sigma^*_{r,n}/\sigma$ is approximately a chi-squared variate over its degrees of freedom with $2/C_{r,n}$ degrees of freedom. Here the authors used the two-moment fit of Patnaik [19] of a weighted chi-square

$$\left[\sigma^*_{r,n} = \sum_{i=1}^{r-1} k_i (X_{i+1,n} - X_{i,n}) \text{ with mean } m = \sigma \text{ and variance } v = C_{r,m}\sigma^2\right]$$

to a chi-square ($2m\sigma^*_{r,n}/v$ with $2m^2/v$ degrees of freedom). Results of Grubbs, Coon and Pearson [6] and Fertig and Mann [5] indicate that under certain conditions this result is applicable also to efficient unbiased linear or maximum-likelihood estimators of Gaussian scale parameters.

Now for either an extreme-value or normal distribution, consider an efficient unbiased linear estimator $\mu^*_{r,n}$ of μ having variance $A_{r,n}\sigma^2$ and covariance $B_{r,n}\sigma^2$ with $\sigma^*_{r,n}$. The estimators $\mu^*_{r,n}$ and $\sigma^*_{r,n}$ are best linear unbiased estimators or efficient simplified approximations thereof. Form the statistic $X^*_{r,n} = \mu^*_{r,n} - (B_{r,n}/C_{r,n})\sigma^*_{r,n}$, which can easily be seen to have covariance $\sigma^2[B - (B/C)C] \equiv 0$ with $\sigma^*_{r,n}$. Let x_P be the $100P$th percentile of the distribution of X, with $P = 1 - R$. It is for x_P, or $\exp(x_P)$, that a lower confidence bound is desired. The parameter x_P is of the form $\mu + z_P\sigma$, where if X is a normal variate, then z_P is the $100P$th percentile of a standard normal distribution with mean zero and variance unity. If X is an extreme-value variate, then, from (1.1), $z_P = \ln\ln[(1-P)^{-1}]$.

If one now forms $X^*_{r,n} - x_P$, as suggested in [18], it can be seen that the expectation E_1 of this difference is given by $E_1 = [(-B_{r,n}/C_{r,n}) - z_P]\sigma$. It can be shown that if P is sufficiently small for a specified r and n, $(X^*_{r,n} - x_P)/E_1$ is, with high probability, a positive random variate. Hence one can combine results of Fertig and Mann [5] and Mann [12], applying to prediction intervals, with previous empirical results of [8, 18] to infer that for appropriate combinations of r, n and p,

$$F_x=(X^*_{r,n}-x_P)/[\sigma^*(-B_{r,n}/C_{r,n}-z_P)], \tag{2.2}$$

has an approximate F distribution. The precision of the approximation is investigated in Section 3.

The number ν_1 of degrees of freedom for the numerator $X^*_{r,n}-x_P$ of F_x can be obtained from Patnaik's [19] two-moment fit: $(X^*_{r,n}-x_P)/m$, with $m=E_1$, is approximately a chi-squared variate over its degrees of freedom, and its number of degrees of freedom is equal to $2m^2/v$, where v is

$$\text{Var}\,(X^*)=(A_{r,n}-B^2_{r,n}/C_{r,n})\sigma^2. \tag{2.3}$$

Therefore, the numbers of degrees of freedom for the approximate F-variate F_x are

$$\nu_1=2(-B_{r,n}/C_{r,n}-z_P)^2/(A_{r,n}-B^2_{r,n}/C_{r,n}) \tag{2.4}$$

and

$$\nu_2=2/C_{r,n}. \tag{2.5}$$

We can now determine, from (2.2), an approximate 100γ percent lower confidence bound for $t_P=\exp(x_P)$ as

$$\exp\{\mu^*_{r,n}-[B_{r,n}/C_{r,n}-\{B_{r,n}/C_{r,n}+z_P\}F_\gamma(\nu_1,\nu_2)]\sigma^*_{r,n}\} \tag{2.6}$$

where $F_\gamma(\nu_1,\nu_2)$ is the 100γth percentile of an F distribution with ν_1 and ν_2 degrees of freedom.

If ν_1 and ν_2 are not integers, which will generally be the case, then one can interpolate in tables of percentiles of the F distribution or use an approximation given by Mann and Grubbs [17].

If $z_P=0$ and $B_{r,n}/C_{r,n}$ is positive (r/n is small), then clearly a lower confidence bound at level γ for $\exp(\mu)$ is obtained by substituting $F_{1-\gamma}(\nu_1,\nu_2)$ for $F_\gamma(\nu_1,\nu_2)>F_{1-\gamma}(\nu_1,\nu_2)$ in (2.6). Using $F_\gamma(\nu_1,\nu_2)$ in (2.6) when $B_{r,n}/C_{r,n}$ is positive will yield an upper confidence bound for $\exp(\mu)$ at level γ.

If best linear invariant estimators $\tilde{\sigma}_{r,n}=\sigma^*_{r,n}/(1+C_{r,n})$ and $\tilde{\mu}_{r,n}=\mu^*-B_{r,n}\tilde{\sigma}_{r,n}$ have been used to estimate μ and σ, then one can use

$$\exp\{\tilde{\mu}_{r,n}+\{B_{r,n}+(1+C_{r,n})[z_P F_\gamma(\nu_1,\nu_2)-\{1-F_\gamma(\nu_1,\nu_2)\}B_{r,n}/C_{r,n}]\}\tilde{\sigma}_{r,n}\} \tag{2.7}$$

to determine an approximate 100γ percent lower confidence bound for t_P. Here values of $A_{r,n}$, $B_{r,n}$ and $C_{r,n}$ can be determined from tabulated values (see [11, 18]) of $E(L\mu)$ and $E(L\sigma)$, the mean squared errors of $\tilde{\mu}/\sigma$ and $\tilde{\sigma}/\sigma$, respectively, and the expected cross product $E(CP)=E[(\tilde{\mu}-\mu)(\tilde{\sigma}-\sigma)]/\sigma^2$. To do this, one uses the following relationships:

$$C_{r,n}=E(L\sigma)/[1-E(L\sigma)] \tag{2.8}$$

$$B_{r,n}=E(CP)/[1-E(L\sigma)] \tag{2.9}$$

and

$$A_{r,n}=E(L\mu)+[E(CP)]^2/[1-E(L\sigma)]. \tag{2.10}$$

For large sample sizes, simplified efficient unbiased linear estimators of μ and σ (see [1], [4] and [15]) can be used in (2.6) with corresponding values of $A_{r,n}$, $B_{r,n}$ and $C_{r,n}\equiv\ell_{r,n}$. Also, as will be shown in Section 3, maximum-likelihood estimators of μ and σ can be substituted for $\tilde{\mu}_{r,n}$ and $\tilde{\sigma}_{r,n}$ in (2.7). In the latter case, values of $A_{r,n}$, $B_{r,n}$ and $C_{r,n}$ corresponding to linear estimators are still appropriate, where available.

3. PRECISION OF THE APPROXIMATION

Several examples are given in [18] showing the excellent agreement of Weibull tolerance bounds based on (2.7) with exact (Monte Carlo) tolerance bounds obtained from tables of [14, 16]. We show now some further examples, among several cases investigated in this present study applying to both Weibull and Gaussian data. And we attempt to draw some guidelines regarding the precision of the approximation.

3.1. Weibull Tolerance Bounds

In the tables below, values of P are those considered in [13], i.e., $P=.01$, .05, .10 and $P=1-e^{-1}$ (corresponding to $z_P=0$, which implies the confidence bound is for $\exp(\mu)$, the Weibull scale parameter, or characteristic life). For each combination of values of P, r and n, the corresponding ν_1 is displayed. Since ν_2 depends only on n and r, it is exhibited only once. Exact table entries $v_{P,\gamma}$ are among those found in [14, 16], and for $\gamma \geq .5$ such that

$$P[x_P > \tilde{\mu}_{r,n} - v_{P,\gamma}\tilde{\sigma}_{r,n}] = \gamma.$$

More generally $v_{P,\gamma}$ is the 100γth percentile of $V_P=(\tilde{\mu}_{r,n}-x_P)/\tilde{\sigma}_{r,n}$. The approximate values of $v_{P,\gamma}$ are calculated from the coefficient of $\tilde{\sigma}_{r,n}$ in (2.7).

Values of $A_{r,n}$, $B_{r,n}$ and $C_{r,n}$ have been obtained by the linear transformations (2.8), (2.9) and (2.10) given in Section 2. The transformations are applied to $E(L\mu)$, $E(L\sigma)$ and $E(CP)$, which are values proportional to expected squared errors and cross products for best linear invariant estimators, tabulated in [11, 12, 18]. The approximation described in [17] has been used in each case for $F_\gamma(\nu_1, \nu_2)$ since ν_1 and ν_2 are, in general, not integers.

Table 1 gives a representation of 3 out of 15 combinations of sample size of n and censoring number r investigated in the present study. The combinations of n and r for the 15 cases were $n=5(5)20$, $r=5(5)n$ and $n=24$, $r=5(5)20$, 24. Approximate values in Table 1 in error by 2 percent or more are bracketed.

Lawless [8] used the F approximation as described in [18], with tables of $A_{r,n}$, $B_{r,n}$ and $C_{r,n}$ applying to simplified efficient unbiased linear estimators, in conjunction with maximum-likelihood estimators. In other words, he substituted maximum-likelihood estimators for simplified versions of best linear invariant estimators in (2.7). He then compared values for obtaining 95 percent tolerance bounds based on (2.7) with exact values computed for maximum-likelihood estimators using numerical integration procedures. The cases Lawless considered involved large values of n, $n=25$, 40 and 60 and values of r/n ranging from .1 to .9. The values of P he used were .05 and .10.

For all of the 20 cases investigated by Lawless for which $r/n \geq .3$, the difference between the 95 percent tolerance bounds based on the F approximation and the exact tolerance bounds based on the numerical integration procedures is within about 1 percent. For $r/n=.1$ or .2 and $P=.05$ the difference is within 3 percent. For very extreme censoring ($r/n=.1$) the approximation (2.7) gives extremely poor results for $P=.10$, apparently because of the small size of ν_1 for these cases.

For the cases investigated in the present study, which clearly involved smaller values of sample size n (and linear, rather than maximum-likelihood estimators), the values of $v_{P,\gamma}$ for obtaining lower confidence bounds on 1st percentiles tended to be within 1 percent of exact (Monte Carlo) values except for samples of sizes 5 and 10. For $P=.05$ and .10 there were more discrepancies

TABLE 1.—*Approximate (from (2.7)) and Exact 100γth Percentiles* $v_{P,\gamma}$ *of* $V_P=(\tilde{\mu}_{r,n}-x)/\tilde{\sigma}_{r,n}$ *for Extreme-Value Data*

Approximate Values from (2.7)				Exact Values from [13]			
$P=1-e^{-1}$	.10	.05	.01	$P=1-e^{-1}$	.10	.05	.01
$n=24, r=5, \nu_2=8.46$							
$\gamma \backslash \nu_1$	23.93	3.50	16.28	79.83	23.93	3.50	16.28	79.83
.02	<−6.704>	<1.691>	<2.140>	3.152	−6.87	0.91	1.90	3.12
.05	<−4.469>	<1.737>	<2.279>	3.441	−4.59	1.36	2.17	3.41
.10	−3.090	<1.799>	<2.433>	3.755	−3.13	1.64	2.38	3.73
.50	<−0.490>	<2.301>	<3.351>	5.575	−0.47	2.45	3.40	5.59
.90	0.585	3.879	5.606	9.943	0.58	3.85	5.64	10.09
.95	0.763	<4.802>	6.808	12.250	0.76	4.50	6.72	12.23
.98	0.929	<6.377>	8.762	15.980	0.92	5.60	8.52	15.75
$n=24, r=10, \nu_2=20.32$								
$\gamma \backslash \nu_1$	12.75	38.92	87.92	271.66	12.75	38.92	87.92	271.66
.02	<−1.720>	<1.552>	2.051	3.126	−1.67	1.47	2.01	3.11
.05	−1.228	<1.673>	2.221	3.408	−1.20	1.63	2.20	3.39
.10	−0.877	1.797	2.393	3.695	−0.87	1.78	2.38	3.69
.50	<−0.060>	2.415	3.242	5.101	−0.08	2.44	3.25	5.10
.90	<0.373>	3.501	4.715	7.532	0.39	3.51	4.73	7.55
.95	<0.454>	3.952	5.324	8.535	0.49	3.94	5.33	8.52
.98	<0.530>	4.575	6.163	9.917	0.60	4.51	6.13	9.94
$n=10, r=10, \nu_2=27.94$								
$\gamma \backslash \nu_1$	1.77	123.14	202.22	453.40	1.77	123.14	202.22	453.40
.02	<−0.320>	<1.257>	1.757	2.864	−0.80	1.21	1.72	2.84
.05	<−0.294>	1.441	1.985	3.195	−0.60	1.42	1.96	3.19
.10	<−0.281>	1.624	2.213	3.526	−0.44	1.62	2.21	3.53
.50	<−0.101>	2.484	3.278	5.071	0.04	2.50	3.29	5.07
.90	<0.679>	3.857	4.974	7.529	0.54	3.86	4.98	7.57
.95	<1.140>	4.390	5.633	8.484	0.71	4.41	5.67	8.57
.98	<1.878>	5.101	6.510	9.753	0.92	5.16	6.65	10.03

larger than 1 percent. For any combination of ν_1 and ν_2, and $\gamma=.75, .90, .95, .98$, the exact and the approximate values of $v_{P,\gamma}$ in both the Lawless study and the present study tend to show excellent agreement (less than about 1 percent error) in the range

$$\nu_2 \geq 5.5,\ 2\nu_2 < \nu_1 < 18\nu_2 - 50. \tag{3.1}$$

(This is similar in spirit to ranges given by Fertig and Mann [5] and Mann [13] applying to precision of approximate prediction intervals for Gaussian and Weibull data, respectively.)

If $P=1-e^{-1}$ so that $x_P=\mu$, there is little discrepancy between the exact and approximate values of $v_{P,\gamma}$, $.02 \leq \gamma \leq .98$, for $r/n \leq .4$, $n \geq 25$. For $r/n \geq .4$, $n \geq 15$, a chi-square approximation discussed in [15, 18, pp. 245–248] can be used.

3.2. Gaussian Tolerance Bounds

For evaluation of the F approximation for Gaussian data, tables of Locks, Alexander and Byars [10] of the noncentral t-distribution were used. In each case considered, of course, $r=n$, and values of $A_{r,n}$, $B_{r,n}$ and $C_{r,n}$ corresponding to unbiased versions of maximum-likelihood estimates were employed. We point out, however, that if alternatively, we had used values corresponding to variances and covariances of best linear unbiased estimators of μ and σ (see Sarhan and Greenberg [21]), ν_1 would not be changed at all and ν_2 would be changed by 1 percent or less.

The noncentral t-variate can be defined by

$$(\overline{X}-\mu-z_P\sigma)/(S/\sqrt{n})$$

with noncentrality parameter $\delta \equiv -z_P\sqrt{n}$ (and $\mu+z_P\sigma=x_P$) and with degrees of freedom $n-1$. Here

$$\overline{X} \equiv \sum_{i=1}^{n} X_i/n$$

is the mean of n normal variatiates with expectation μ, and

$$S \equiv \sum_{i=1}^{n} (X_i-\overline{X})^2/(n-1)$$

has expectation $c_n\sigma$, with $c_n=\sqrt{2/(n-1)}\,\Gamma(n/2)/\Gamma[(n-1)/2]$. Then since $\overline{X}$ and S are independent, one can form an approximate F-statistic as

$$(\overline{X}-x_P)/(-z_PS/c_n) \tag{3.2}$$

with

$$\nu_1=2nz_P^2 \tag{3.3}$$

and

$$\nu_2=2c_n^2/(1-c_n^2). \tag{3.4}$$

Then, percentiles of the distribution of the approximate F-statistic (3.2) can be used to approximate percentiles of the noncentral t-distribution by multiplying the former by $-z_p\sqrt{n}/c_n$. In Table 2, exact 100γth percentiles of the noncentral t-distribution from tables of [9] for $\gamma=.005, .01, .05, .10, .25, .50, .75, .90, .95, .99$, and $-z_P=1.0(.50)3.0$ are tabulated for a specified $n-1$, with $n-1=10$

and 35. For each value of $n-1$ a tabulation of corresponding percentiles based on the F approximation (3.2), (3.3) and (3.4) is exhibited. Approximate values in error by more than two percent are bracketed.

It can be seen from Table 2 that the approximation is excellent for $\gamma=.5$ or more (the range of present interest) for $n\geq 9$ (or $\nu_2\geq 27.1$) when one is concerned with 0.14th through 16th percentiles of normal and lognormal distributions. The tabulated results, of course, apply also to 84th through 99.86th distribution percentiles when the sign of z_P is changed. The ranges over which the error in the F approximation is about 2 percent or less are given roughly by

$$\gamma=.99:\nu_2\geq 27,\ .4\nu_2+10<\nu_1<5.2\nu_2-90 \tag{3.5}$$

$$\gamma=.95:\nu_2\geq 27,\ .4\nu_2+10<\nu_1<5.9\nu_2-22 \tag{3.6}$$

$$\gamma=.90:\nu_2\geq 27,\ .4\nu_2+10<\nu_1 \tag{3.7}$$

$$\gamma=.50,\ .75:\nu_2\geq 27,\ .7\nu_2+15<\nu_1 \tag{3.8}$$

TABLE 2.—*Approximate (from (2.7)) and Exact Percentiles of the Noncentral t-distribution with Noncentrality Parameter* $-\sqrt{n}z_P$

Approximate Values from (2.7); $n-1=10$, $\nu_2=39.1$						Exact Values from [9]; $n-1=10$, $\nu_2=39.1$					
$-Z_P$	1.00	1.50	2.00	2.50	3.00	$-Z_P$	1.00	1.50	2.00	2.50	3.00
$\gamma \backslash \nu_1$	22.0	49.4	87.9	137.5	198.0	$\gamma \backslash \nu_1$	22.0	49.4	87.9	137.5	198.0
.005	<1.193>	<2.348>	3.482	4.585	5.666	.005	0.725	2.156	3.428	4.618	5.761
.01	<1.326>	<2.533>	3.713	4.863	5.990	.01	0.961	2.385	3.674	4.892	6.069
.05	<2.032>	<3.472>	4.430	5.718	6.987	.05	1.608	3.053	4.414	5.729	7.017
.10	<2.418>	<3.965>	4.872	6.243	7.598	.10	1.966	3.443	4.859	6.240	7.601
.25	<2.582>	4.171	5.723	7.254	8.772	.25	2.606	4.170	5.705	7.223	8.730
.50	<3.247>	5.116	6.864	8.605	10.342	.50	3.412	5.127	6.844	8.562	10.279
.75	<4.322>	6.289	8.272	10.267	12.270	.75	4.372	6.308	8.276	10.261	12.256
.90	5.435	7.600	9.836	12.110	14.406	.90	5.434	7.647	9.918	12.224	14.550
.95	6.232	8.526	10.936	<13.404>	<15.905>	.95	6.193	8.618	11.119	13.665	16.238
.99	8.067	<10.628>	<13.423>	<16.327>	<19.289>	.99	7.968	10.916	13.979	17.109	
Approximate Values from (2.7); $n-1=35$, $\nu_2=138.9$						Exact Values from [9]; $n-1=35$, $\nu_2=138.9$					
$-Z_P$	1.00	1.50	2.00	2.50	3.00	$-Z_P$	1.00	1.50	2.00	2.50	3.00
$\gamma \backslash \nu_1$	72.0	162.0	288.0	450.0	648.0	$\gamma \backslash \nu_1$	72.0	162.0	288.0	450.0	648.0
.005	<3.490>	5.958	8.373	10.745	13.088	.005	3.255	5.873	8.368	10.794	13.178
.01	<3.684>	6.206	8.674	11.101	13.501	.01	3.499	6.139	8.671	11.141	13.574
.05	4.602	6.934	9.555	12.142	14.707	.05	4.188	6.908	9.554	12.158	14.737
.10	5.047	7.356	10.064	12.742	15.402	.10	4.572	7.345	10.062	12.747	15.412
.25	5.228	8.123	10.984	13.827	16.658	.25	5.244	8.124	10.977	13.812	16.638
.50	6.011	9.069	12.115	15.157	18.197	.50	6.048	9.076	12.106	15.136	18.166
.75	6.903	10.131	13.379	16.642	19.914	.75	6.928	10.138	13.377	16.634	19.901
.90	7.815	11.201	14.649	18.132	21.636	.90	7.799	11.206	14.667	18.161	21.674
.95	8.413	11.898	15.473	19.098	22.751	.95	8.364	11.906	15.517	19.169	22.847
.99	9.658	13.334	17.169	21.083	25.043	.99	9.528	13.362	17.294	21.285	25.312

For $.50<\gamma<.90$, the upper limit on $v_1=2nz_p^2$ corresponds to the extreme upper or lower tail of a normal or lognormal distribution.

These ranges should be applicable even when samples are censored. Thus, the approximation (2.6) can be used with best linear unbiased estimates of normal parameters to obtain lognormal tolerance bounds when ν_1 and ν_2 given by (2.4) and (2.5) fall within one of the various ranges specified by (3.5) through (3.8). Also, work has begun by this author to generate constants for use in obtaining simplified linear estimates from large censored normal and lognormal samples. It should then be possible to use these in conjunction with the approximation (2.6).

REFERENCES

[1] Bain, L. J., "Inferences Based on Censored Sampling from the Weibull or Extreme-Value Distribution," Technometrics, *14*, 693–702 (1962).

[2] Billman, B. R., C. L. Antle, and L. J. Bain, "Statistical Inference from Censored Weibull Samples," Technometrics, *14*, 831–840 (1972).

[3] Box, G. E. P., "Some Theorems on Quadratic Forms Applied in the Study of Variance Problems, I. Effect of Inequality of Variance in the One-Way Classification," Ann. Math. Statist., *65*, 290–302 (1954).

[4] Englehardt, M. and L. J. Bain, "Some Complete and Censored Sampling Results for the Weibull or Extreme-Value Distribution," Technometrics, *15*, 541–549 (1973).

[5] Fertig, K. W. and N. R. Mann, "A New Approach to the Determination of Exact and Approximate One-Sided Prediction Intervals for Normal and Lognormal Distribtuions, with Tables," in *Reliability and Fault-Tree Analysis*, R. Barlow, Ed., SIAM Series in Applied Mathematics (1974).

[6] Grubbs, F. E., H. J. Coon, and E. S. Pearson, "On the Use of Patnaik Type Approximations to the Range in Significance Tests," Biometrika, *53*, 248–252 (1966).

[7] Lawless, J. F., "Conditional Versus Unconditional Confidence Intervals for the Parameters of the Weibull Distribution," J. Amer. Statist. Assoc., *68*, 665–669 (1973).

[8] Lawless, J. F., "Construction of Tolerance Bounds for the Extreme-Value and Weibull Distributions," Technometrics, *17*, 255–261 (1975).

[9] Lehman, E. L., *Testing Statistical Hypotheses* (John Wiley, New York 1959).

[10] Locks, M. O., M. J. Alexander and B. J. Byars, "New Tables of the Noncentral t Distribution," Aerospace Research Laboratories Report ARL 63–19, Aerospace Research Laboratories, Wright-Patterson Air Force Base, Ohio (1963).

[11] Mann, N. R., "Results on Location and Scale Parameter Estimation with Application to the Extreme-Value Distribution," Aerospace Research Laboratories Report ARL 67–0023; Office of Aerospace Research, U.S. Air Force, Wright-Patterson Air Force Base, Ohio (1967).

[12] Mann, N. R., "Tables for Obtaining the Best Linear Invariant Estimates of Parameters of the Weibull Distribution," Technometrics, *9*, 629–645 (1967).

[13] Mann, N. R., "Warranty Periods for Production Lots Based on Fatigue-Test Data," Engineering Fracture Mechanics, *8*, 123–130 (1976).

[14] Mann, N. R. and K. W. Fertig, "Tables for Obtaining Confidence Bounds and Tolerance Bounds Based on Best Linear Invariant Estimates of Parameters of the Extreme-Value Distribution," Technometrics, *15*, 87–101 (1973).

[15] Mann, N. R. and K. W. Fertig, "Simplified Efficient Point and Interval Estimators for Weibull Parameters," Technometrics, *17*, 361–368 (1975).

[16] Mann, N. R., K. W. Fertig, and E. M. Scheuer, "Confidence and Tolerance Bounds and a New Goodness-of-Fit Test for Two-Parameters Weibull or Extreme-Value Distributions with Tables for Censored Samples of Size 3(1)25," Aerospace Research Laboratories Report ARL 71–0077, Office of Aerospace Research, United States Air Force, Wright-Patterson Air Force Base, Ohio (1971).

[17] Mann, N. R. and F. E. Grubbs, "Simple, Efficient Closed-Form Approximations for Beta Percentiles, Exponential Prediction Intervals and Confidence Bounds on Exponential and Binominal Parameters," J. Amer. Statist. Assoc., *66*, 654–661 (1974).

[18] Mann, N. R., R. E. Schafer, and N. D. Singpurwalla, *Methods for the Statistical Analysis of Reliability and Life Data* (John Wiley, New York 1974).

[19] Patnaik, P. B., "The Non-Central χ^2 and F Distributions and Their Applications," Biometrika, *36*, 202–232 (1949).

[20] Pyke, R., "Spacings," J. Royal Statist. Soc. B, *27*, 395–449 (1965).

[21] Sarhan, A. E. and B. G. Greenberg, *Contributions to Order Statistics* (John Wiley 1962).

[22] Thoman, D. R., L. J. Bain, and C. E. Antle, "Inferences on the Parameters of the Weibull Distribution," Technometrics, *11*, 445–460 (1969).

[23] van Montfort, M. A. J., "On Testing that the Distribution of Extremes is of Type I when Type II is the Alternative," Journal of Hydrology, *11*, 421–427 (1970).

A NOTE ON A CONFIDENCE INTERVAL FOR AN INTERCLASS MEAN

T. Jayachandran*

Naval Postgraduate School
Monterey, California

ABSTRACT

An exact confidence interval for an interclass mean, that is, the mean of a composite sample made up several subsamples of unequal sizes n_i is presented.

1. INTRODUCTION

Suppose X_{ij}, $j=1, 2, \ldots, n_i$; $i=1, 2, \ldots, k$ is a composite sample of size

$$N=\sum_{i=1}^{k} n_i$$

that is comprised of k subsamples of sizes n_i. The i^{th} subsample is a random sample from a normal distribution $N(a_i,\sigma^2)$ and the a_i's, $i=1, 2, \ldots, k$ are assumed to be independent and identically distributed $(i\cdot i\cdot d)$ as $N(\mu, \sigma_a^2)$. μ is sometimes known as the interclass mean and the problem considered in this note is the construction of a confidence interval for μ. Some practical uses of such an interval are discussed in a paper by Long [1]. For example, the composite sample may be measurements on a characteristic of the output of a factory made on different days. μ would correspond to the true mean value of the characteristic being studied.

The construction of a confidence interval for μ is straightforward if all the n_i are equal. For unequal subsample sizes Long [1] obtained an approximate interval that he shows to be reasonably accurate. The procedure proposed in this paper leads to an exact confidence interval for μ in the case of unequal n_i.

2. PRELIMINARIES

The observations X_{ij}, $j=1, 2, \ldots, n_i$; $i=1, 2, \ldots, k$ may be assumed to satisfy the variance components model

$$Y_{ij}=a_i+e_{ij},\ j=1, 2, \ldots, n;\ i=1, 2, \ldots, k \tag{2.1}$$

*This research was supported by the Office of Naval Research as part of the Foundation Research Program at the Naval Postgraduate School.

where the a_i's are $i{\cdot}i{\cdot}d.$ $N(\mu\sigma_a^2)$, the e_{ij}'s are $i{\cdot}i{\cdot}d.$ $N(0, \sigma^2)$ and the a_i's and e_{ij}'s are mutually independent. Let

$$\overline{X}_i=\sum_{j=1}^{n_i}\frac{X_{ij}}{n_i} \text{ and } \overline{X}=\sum_{i=1}^{k}\frac{\overline{X}_i}{k}$$

and

$$S^2=\sum_{i=1}^{k}\frac{(\overline{X}_i-\overline{X})^2}{k-1}$$

Then, $\overline{X}_i$, $i=1, 2, \ldots, k$ are independent and normally distributed

$$N\left(\mu, \sigma_a^2+\frac{\sigma^2}{n_i}\right).$$

If $n_i=n$ for all i the $\overline{X}_i$ will constitute a random sample from a normal distribution

$$N\left(\mu, \sigma_a^2+\frac{\sigma^2}{n}\right).$$

Thus,

$$T=\frac{k^{1/2}(\overline{X}-\mu)}{S} \tag{2.2}$$

has a student's t distribution with $(k\text{-}1)$ degrees of freedom, and this leads to a confidence interval or μ.

For the case where the n_i are not all equal, Long [1] obtained an approximate confidence interval for μ by treating T as a student's t variable. On the basis of a study of the exact distribution fof T for $k=2$, 3 and an examination of the moments of T for large values of k, Long [1] concludes that the t approximation is fairly reliable. He also points out that the approximation is likely to go wrong if there are wide variations in the subsample sizes n_i or if σ^2 is large relative to σ_a^2.

3. EXACT CONFIDENCE INTERVAL WITH UNEQUAL SUBSAMPLE SIZES

For unequal subsample sizes, the $\overline{X}_i$ are not identically distributed since the variance of $\overline{X}_i$ is

$$\sigma_a^2+\frac{\sigma^2}{n_i},$$

even though all of them have the same mean μ. Thus, the construction of a student's t variable, based on $\overline{X}_i$ only, independent of the nuisance parameters $\sigma_a^2+\sigma^2/n_i$, is not possible. To get around this difficulty let

$$Z_i=C_{i1}X_{i1}+C_{i2}\overline{X}_i \qquad i=1, 2, \ldots k \tag{3.1}$$

where

$$C_{i1}=\left(\frac{n_iC-1}{n_i-1}\right)^{1/2}$$

$$C_{i2}=1-C_{i1}$$

and

$$C=\frac{1}{\min n_i}.$$

Then, as shown in the appendix,

$$E(Z_i)=\mu$$

$$V(Z_i)=\sigma_a^2+C\sigma^2=\sigma_a^2+\frac{\sigma^2}{\min n_i}$$

If

$$\overline{Z}=\frac{1}{k}\sum_{i=1}^{k} Z_i$$

and

$$S_Z^2=\frac{1}{k-1}\sum_{i=1}^{k}(Z_i-\overline{Z})^2$$

then

$$T=\frac{k^{1/2}(\overline{Z}-\mu)}{S_Z} \tag{3.2}$$

has a student's t-distribution with $k-1$ degrees of freedom. An exact confidence interval for μ is now obtainable in the usual way.

4. DISCUSSION

The exact procedure proposed in Section 3 is somewhat ad hoc in nature in the sense that the first observation X_{i1} in each subsample plays a prominent role. To avoid any systematic bias that may creep in, it would be preferable to randomly permute the observations in each of the k subsamples before applying the procedure. If all the subsample sizes n_i are equal to n, the statistic (3.2) reduces to the usual T statistic (2.2).

5. APPENDIX

Under the assumptions of the variance components model (2.1)

$$E(X_{ij})=\mu \qquad j=1, 2, \ldots, n_i;\ i=1, 2, \ldots, k$$

$$\text{Var } (X_{ij})=\sigma^2+\sigma_a^2$$

and

$$\text{Cov } (X_{ij}, X_{i'j'})=\begin{cases}0 & i\neq i'\\ \sigma_a^2 & i=i',\ j\neq j'\end{cases}$$

t follows

$$E(\overline{X}_i)=\mu$$

$$V(\overline{X}_i)=\sigma_a^2+\frac{\sigma^2}{n_i} \qquad i=1, 2, \ldots, k$$

nd

$$\text{Cov } (X_{ij}, \overline{X}_i)=\sigma_a^2+\frac{\sigma^2}{n_i} \qquad j=1, 2, \ldots, n_i$$

Therefore, if

$$Z_i = C_{i1}X_{i1} + C_{i2}\overline{X}_i$$

then

$$E(Z_i) = (C_{i1} + C_{i2})\mu = \mu$$

and

$$V(Z_i) = C_{i1}^2 V(X_{i1}) + C_{i2}^2 V(\overline{X}_i) + 2C_{i1}C_{i2}\,\mathrm{Cov}\,(X_{i1}, \overline{X}_i)$$

$$= C_{i1}^2(\sigma^2 + \sigma_a^2) + (C_i^2 + 2C_{i1}C_{i2})\left(\sigma_a^2 + \frac{\sigma^2}{n_i}\right).$$

substitution of

$$C_{i1} = \left(\frac{n_i C - 1}{n_i - 1}\right)^{1/2},\ C_{i2} = 1 - C_{i1} \text{ and } C = \frac{1}{\min n_i}$$

results in

$$E(Z_i) = \mu \text{ and } V(Z_i) = \sigma_a^2 + \frac{\sigma^2}{\min n_i}.$$

REFERENCES

1. Long, W. M., "Estimation Problems When a Simple Type of Heterogeneity is Present in the Sample," Biometrika, *38*, 90–101 (1951).

*U.S. GOVERNMENT PRINTING OFFICE: 1977240-830/1 1-3

INFORMATION FOR CONTRIBUTORS

The NAVAL RESEARCH LOGISTICS QUARTERLY is devoted to the dissemination of scientific information in logistics and will publish research and expository papers, including those in certain areas of mathematics, statistics, and economics, relevant to the over-all effort to improve the efficiency and effectiveness of logistics operations.

Manuscripts and other items for publication should be sent to The Managing Editor, NAVAL RESEARCH LOGISTICS QUARTERLY, Office of Naval Research, Arlington, Va. 22217. Each manuscript which is considered to be suitable material for the QUARTERLY is sent to one or more referees.

Manuscripts submitted for publication should be typewritten, double-spaced, and the author should retain a copy. Refereeing may be expedited if an extra copy of the manuscript is submitted with the original.

A short abstract (not over 400 words) should accompany each manuscript. This will appear at the head of the published paper in the QUARTERLY.

There is no authorization for compensation to authors for papers which have been accepted for publication. Authors will receive 250 reprints of their published papers.

Readers are invited to submit to the Managing Editor items of general interest in the field of logistics, for possible publication in the NEWS AND MEMORANDA or NOTES sections of the QUARTERLY.

NAVAL RESEARCH LOGISTICS QUARTERLY

MARCH 1977
VOL. 24, NO. 1
NAVSO P-1278

CONTENTS

ARTICLES

		Page
A Two-Echelon Inventory Model with Purchases, Dispositions, Shipments, Returns and Transshipments	B. HOADLEY D. P. HEYMAN	1
Optimal Reject Allowance with Constant Marginal Production Efficiency	A. BEJA	21
A Chance-Constrained Distribution Problem	R. M. REESE A. C. STEDRY	35
Elements of a Theory in Non-Convex Programming	C. BURDET	47
Convex and Polaroid Extensions	C. BURDET	67
A Cutting Plane Algorithm for the Bilinear Programming Problem	H. VAISH	83
The Effect of Correlated Exponential Service Times on Single Server Tandem Queues	C. R. MITCHELL A. S. PAULSON C. A. BESWICK	95
Single-Lane Bridge Serving Two-Lane Traffic	Z. ESHCOLI I. ADIRI	113
Optimal Control for Multi-Servers Queueing Systems under Periodic Review	C. C. HUANG S. L. BRUMELLE K. SAWAKI I. VERTINSKY	127
Cyclical Job Sequencing on Multiple Sets of Identical Machines	H. I. STERN E. P. RODRIGUEZ M. L. UTTER	137
Johnson's Approximate Method for the 3 X n Job Shop Problem	W. SZWARC G. K. HUTCHINSON	153
A Convex Property of an Ordered Flow Shop Sequencing Problem	S. S. PANWALKAR A. W. KHAN	159
A Manpower Planning/Capital Budgeting Model (MAPCAB)	R. H. CLARK R. A. COMERFORD	163
An F Approximation for Two-Parameter Weibull and Lognormal Tolerance Bounds Based on Possibly Censored Data	N. R. MANN	187
A Note on a Confidence Interval for an Interclass Mean	T. JAYACHANDRAN	197

OFFICE OF NAVAL RESEARCH
Arlington, Va. 22217